Transport Policy 133 (2023) 1–14

Contents lists available at ScienceDirect

Transport Policy

journal homepage: www.elsevier.com/locate/tranpol

Analysis of gender-specific bicycle route choices using revealed preference surveys based on GPS traces

Federico Rupi [a,*], Marzia Freo [b], Cristian Poliziani [c], Maria Nadia Postorino [a], Joerg Schweizer [a]

[a] Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Italy
[b] European Commission, Joint Research Center (JRC), Italy
[c] Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA, 94720, USA

ARTICLE INFO

Keywords:
User behaviour
Path choices
Bike gender gap
GPS data
Discrimination factor

ABSTRACT

Bike facility features in urban transport systems are one of the most important elements for encouraging user choices regarding sustainable transport modes. The process of designing the bikeway does involve biker perception but the act of designing does not often rely on this perception. In order to identify whether gender differences exist for bike route choices, the actual choices made by bikers - both male and female - have been detected by means of GPS data, with the pathways characteristics being known. Detected route choices have been analyzed using the Oaxaca-Blinder decomposition method (Blinder, 1973; Oaxaca, 1973), which provides a possible explanation for differences in gender-specific route attributes that male and female cyclists experience under similar conditions. The results show that differences between female and male cyclists exist in terms of the ease of use of the pathways and related choices. Some analyses regarding age classes have also shown that gender differences tend to be less relevant with increasing age, thus suggesting that more-experienced female cyclists make choices similar to those of their male counterparts.

1. Introduction

In the last several decades, great attention has been given to detecting, analysing, and simulating bicycle volume flows in urban areas. Data availability is crucial to developing research and to support decision-makers for evidence-based bicycle solutions (Broach et al., 2012). In addition, there is a need to identify locations that lack baseline data for monitoring changes associated with new policies or infrastructure.

In recent years, the increasing number of mobile apps used to track individual bicycling has contributed to record large amounts of geo-referenced bicycle trips (Nelson et al., 2021) and has made it possible to associate user's features, mainly gender and age, to their route choices.

With regard to gender as a factor, several researches have been highlighted the different exigencies of women regarding transit systems (Arabikhan et al., 2016) and that female cyclists are averse to cycling routes with high levels of vehicular traffic and longer distances (Krizek et al., 2005). For example, in a Dutch study, Heinen et al. (2013) showed that for commuter bicycle trips, women are more distance-sensitive than men. The imbalance between men's and women's bicycle use, also referred to as the "bike gender gap," has been frequently reported in countries and cities with low bicycling rates. According to several studies, women are less likely than men to bicycle in most places in the United States, sometimes by a ratio of 3:1 (Dill et al., 2015; Schoner et al., 2015; Akar and Clifton, 2009; Akar et al., 2013; Krizek et al., 2005; Pucher et al., 2011a,b). This conclusion has also been found in studies in Australia (Garrard et al., 2008), Canada (Winters et al., 2011), and New Zealand (Shaw et al., 2020). Many other studies confirm that women are less likely to bicycle compared to men (Emond et al., 2009; Krizek et al., 2005; Moudon et al., 2005; Sener et al., 2009; Stinson and Bhat 2004; Winters et al., 2007). Pucher and Buehler's overview (2008) showed that in Australia, the United Kingdom, the United States, and Canada, the participation rate of women is 30% or less. Winters and Zanotto (2017) showed that in cities in North America, Australia, New Zealand, and the United Kingdom, women are under-represented in the cycling population. In the United States, only 0.6% of the population 16 years of age and older uses bicycles for transportation, and only 28% are female (US Census U.S. Census Bureau, 2019). In Australia, between 17% and 25% of bicycle commuting trips are made by women (Heesch et al., 2012), with absolute numbers being much less relevant in countries with a low (approximately 5%) cycling mode share (Pucher and Buehler,

* Corresponding author.
E-mail address: federico.rupi@unibo.it (F. Rupi).

https://doi.org/10.1016/j.tranpol.2023.01.001
Received 24 February 2022; Received in revised form 2 January 2023; Accepted 4 January 2023
Available online 7 January 2023
0967-070X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

2008). Based on a recent study carried out by the Institute for Transportation and Development Policy (ITDP, 2022), a critical gender gap exists in cities of Brazil, such as Rio de Janeiro, where women range between 2.4% and 10.9% of all cyclists. In addition, in Delhi, India, where 21% of trips are made on bicycles, women constituted only 2% of those riders (ITDP, 2022).

Conversely, women bicycle as much as, if not more than, men in northern European countries such as Denmark, Germany, and the Netherlands across all ages and trip purposes (Pucher and Buehler, 2008). According to several studies (Buehler et al., 2011; Pucher et al., 2011; Emond et al., 2009; Akar et al., 2013), 45% of all bike trips in Denmark, 49% in Germany, and 55% in the Netherlands are made by women. These same researchers have found that communities with higher levels of bicycling tend to have a higher ratio of female to male cyclists.

In Italy, according to the Eurobarometer survey (Special Eurobarometer 422, 2014), only 6% of the respondents reported bicycles as their most frequently used mode of transportation, lower than the average rate in Europe (8%). The figures provided by the Italian National Institute of Statistics (ISTAT) show that only 4.2% of Italian employees commute to work by bicycle, of which 48% are female (ISTAT, 2017; ISTAT, 2020). Although these figures show low rates of bicycling, the population of cyclists in Italy is gender balanced as in Northern Europe countries. In Bologna (a medium-size city in Northern Italy), approximately 7% of the population uses bicycles and, as revealed by the survey data carried out along the Bologna cycle network (Municipality of Bologna, 2020), 45% of the cyclists are female.

To understand how preferences and behaviour of cyclists vary by gender, many studies have focused on the use of stated preferences data, although the use of stated preferences data could have weaknesses (Cherchi and Hensher, 2015). However, there is a lack of studies examining actual choices made by cyclists (revealed preferences), and it remains unclear if and how male and female cyclists choose routes in a mature, gender-balanced cycling environment, as is the case in many European countries (see section 2).

To improve knowledge in this field, this work explores and quantifies potential differences between male and female cyclists and attempts to address the question of whether and how bike route choices and facility preferences are linked to gender in order to provide information for planning and suitably designing bike networks. To this aim, this study is based on commuter cyclists' route choices based on revealed preference data (particularly GPS data) as recorded by the cyclists themselves.

The premise for the investigation of potential gender-related differences is that cyclists departing from a given origin choose a route among a set of available alternatives, based on specific route attributes, to reach a destination. Gender discrimination is assumed to intervene if male and female cyclists, all other individual conditions being equal, choose routes with systematically different attributes to travel the same origin-destination (OD) pair. In terms of route choices, any decision implies a range of observed attributes. The decomposition approach, proposed by Oaxaca and Blinder), is hereby adopted to estimate the role of gender differential in observed attributes depending only on gender discrimination.

The goal of this study is to investigate whether route choices and their attributes are associated with gender discrimination or, in other words, i) if female cyclists choose routes with systematically different attributes as compared to male cyclists and ii) whether the routes chosen by female cyclists are systematically less or more difficult to ride.

The paper is structured as follows: Section 2 focuses on the main findings from existing literature and identifies the addressed research question. Section 3 describes the datasets used in this study to explore cyclists' differences by gender, if any. Section 4 presents the methodological approach used. Section 5 discusses the results and the policy implications, and Section 6 concludes.

2. Literature review

Constructed environment-related variables might affect cyclists' route choices (see Schweizer et al., 2014). In particular, differences might exist in facility preferences related to intersections with or without traffic lights, bicycle lanes, the condition of said lanes, lane organization (mixed traffic or separated lanes), etc. However, there are few studies evaluating how the built environment along the route affects bicycle users' perceptions (Echiburú et al., 2021). Moreover, a small segment of the literature focuses specifically on the gender balance of cyclists (AitBihiOuali and Klingen, 2022).

Some studies have shown clear gender differences, particularly those regarding the relationship between infrastructural improvements and cycling growth rates, with safer infrastructure being more important for female cyclists. Such differences have been explained by assuming higher levels of traffic risk aversion and lower levels of bicycle confidence among female cyclists (Garrard et al., 2008). Safety risks (actual and perceived) associated with cycling (encompassing safe cycleway, road safety, and secure bicycle parking) appear to be a significant deterrent to female cyclists. Male cyclists often seem to be less affected by poor cycling facilities (Garrard, 2003).

Baker (2009) states that the behaviors of female cyclists are considered an indicator of an environment's level of cyclist safety. From this perspective, higher percentages of female riders would indicate safer infrastructure for bicycling. Indeed, Garrard et al. (2008) highlight that if one wants to measure the bicycle friendliness of an urban environment, the ratio of female to male bicyclists would be a good indicator. AitBihiOuali and Klingen (2022) also found that dedicated cycling infrastructure increases women's cycling participation between 4% and 6%.

The arising question is then: are women less risk-prone than men? There is a significant body of literature documenting the fact that women tend to be more risk averse than men and perceive negative consequences of sharing roads with vehicular traffic more than men (Weber et al., 2002; Harris et al., 2006). This suggests lower rates of facility use in such conditions and longer routes with separated bicycle lanes are necessary to increase safety levels (Krizek et al., 2005). This behaviour has been associated with the characteristics of their trips (multipurpose and/or encumbered trips). Similar findings in using safer off-road paths have been described in Australia and China (Heesch et al., 2012; Lusk et al., 2014).

BlaisWeber (2001) argue that, although women appear to be more risk averse, this may be the result of gender differences in risk perceptions rather than the result of more conservative attitudes towards risk-taking. Matsuda et al. (2000) showed that among a population of high school students 16 years of age and older, female students had different perceptions of risk and tended to avoid risky practices while bicycling compared to male students.

The presence of gender differences in traffic risk aversion was confirmed by Jensen et al. (2019), who, using crowd sourced data, found that female cyclists dislike cycling in the wrong direction, perceiving the route as 128% longer, whereas this perception was 101.7% for male cyclists. The core hypothesis regarding gender differences in traffic risk aversion was not fully confirmed by Carroll et al. (2020), who found that gender differences regarding the positive effect of "safe" bicycle lanes in Dublin were not statistically significant.

Garrard et al. (2008), among others, investigated whether female commuter cyclists in Melbourne, Australia, were more likely to use bicycle routes that provide separation from vehicular traffic. They found that, consistent with gender differences in risk aversion, female commuter cyclists preferred to use routes with maximum separation from motorized traffic: the provision of on-road lanes on busy arterial roads may not offer the level of separation from vehicular traffic needed to attract increased numbers of female commuter cyclists. Heesch et al. (2012), by an online survey about cycling patterns carried out in Queensland, Australia, found that male cyclists were more likely to cycle longer and use on-road paths than female cyclists. These findings have been confirmed in a report on seven cities in the United Kingdom (Sustrans, 2018).

Emond et al. (2009) found that both male and female cyclists become less comfortable as the size of the street increases and bike lanes are absent, but the decrease in comfort is more significant for female than male cyclists. They found that male and female cyclists show a significant preference for bicycle facilities that separate them from vehicular traffic, even if the female cyclists are less comfortable on off-street paths than male cyclists. This behaviour could be related to a concern over personal security on segregated and potentially less visible facilities, suggesting a possible trade-off between safety from traffic and safety from attacks.

Based on an online survey, Twaddle et al. (2010) distinguished among potential cyclists and current cyclists with varying degrees of bicycle commuting frequency. They suggested that potential female cyclists might perceive said factors differently and that there may also exist more barriers to cycling than for male cyclists; therefore, they are more concerned than male cyclists about safety and security issues associated with cycling. However, there was no significant difference by gender in the selection of any of the on-route improvements (bicycle lanes, more-direct routes, and more bicycle paths); men and women have similar facility preferences, and the largest percentage of possible cyclists indicated a desire for bicycle lanes. This research contradicts previous works that suggested women prefer to be separated from vehicular traffic (Garrard et al., 2008; Krizek et al., 2005; Nelson and Allen, 1997; Tilahun et al., 2007; Dill and Gliebe, 2008). Moreover, according to Twaddle et al. (2010), for current cyclists, the most requested on-route improvement is the desire for more bicycle lanes, but female cyclists do not appear to have a strong preference for off-road bicycle paths and have an equally strong desire for bicycle lanes as do male cyclists. However, a more recent study (Winters and Zanotto, 2017) has found that women cyclists chose their routes based on facilities with high level of separation from motor vehicles. This conclusion was confirmed by Aldred et al. (2017), who found that female cyclists reported stronger preferences than men for greater separation from vehicular traffic.

Akar et al. (2013) found that cyclists give great importance to travel time and that decreasing travel time will increase the probability of bicycling, which significantly affects female cyclists, as confirmed by Abasahl et al. (2018). Based on the results of their study, feeling safe is significantly and positively associated with bicycling choice, and male and female cyclists have significant differences in this attitude, such that female cyclists are much more concerned about safety (e.g., vehicular traffic and a lack of bicycle lanes, paths, and trails) compared to male cyclists. According to the authors, the reason for this result might be the fact that a higher percentage of women identified themselves as “beginner cyclists'' who are not comfortable with riding with regular traffic. These results were confirmed by Prati et al. (2019), who argue that gender differences towards cycling are not particularly evident among more-experienced cyclists.

By a survey conducted in Hangzou, China, Lusk et al. (2014) found that the difference between male and female cyclist preferences in using vehicular roads was statistically significant, with female cyclists preferring not to use roads used by vehicles. Similarly, the gender-related difference in preference for cycling on cycle tracks was statistically significant, with female cyclists preferring to use them.

Echiburú et al. (2021) argue that the presence of cycle paths and trip length do not have the same effects on male and female cyclists: female cyclists are more sensitive to longer distances than male cyclists, and female cyclists appear to be less satisfied by cycle paths in the route, perhaps because of more competitive and risky manoeuvres among male cyclists on this type of infrastructure.

Grudgings et al. (2021) studied the influence of determinants' interactions with gender or age on cycling behaviours in England and Wales. In this work, they found that determinants associated with physical effort (hilliness and distance) and traffic (traffic density and cycle lanes) are more important in older age groups of both men and women. The analysis employed 17 determinants of commuter cycling mode sharing to describe the utility of the local environment for cycling. Cycling levels are strikingly more sensitive to gender than to age. Female cyclists tend to require a higher threshold of utility, which combines the effect of both place-based and population-based determinants, to start cycling than do male cyclists; in higher-utility environments, gender differences are almost non-existent. Differences in cycling rates by age and gender diminish in more-supportive cycling environments. Therefore, there do not appear to be strong benefits in targeting cycling interventions by both age and gender.

To summarize (Table 1), most studies indicate that female cyclists have stronger preference for safe, protected bikeways compared to male cyclists, but this is primarily the case if said female cyclists are beginners. Some findings indicate that female cyclists would only begin cycling if there was a safe bike network available, whereas in cities with a long-established network of bikeways, the entire cyclist population is

Table 1
Summary of past research.

Contribution	Finding	Gender difference in cycling behaviors
Nelson and Allen (1997)	Higher levels of traffic risk aversion and lower levels of bicycle confidence among female cyclists	Significant
Matsuda et al. (2000)	Different gender perceptions of risk among high school students 16 years of age or older	Significant
BlaisWeber, 2001	Gender differences in risk perceptions	Significant
Weber et al., 2002; Harris et al. (2006)	Greater female concern for sharing roads with vehicular traffic	Significant
Garrard (2003)	Greater female concern for poor cycling facilities	Significant
Krizek et al. (2005); Tilahun et al. (2007)	Greater female preference for separated bicycle lanes	Significant
Garrard et al. (2008), Dill and Gliebe (2008)	Greater female preference for bicycle-friendly urban environments and for routes with maximum separation from vehicular traffic	Significant
Baker (2009)	Greater female preference for safer cycling infrastructure	Significant
Heesch et al. (2012); Lusk et al. (2014); Sustrans (2018)	Greater female preference for using safer off-road paths	Significant
Emond et al. (2009); Echiburú et al. (2021)	Lower level of comfort among female cyclists using off-street paths (personal security)	Significant
Winters and Zanotto (2017); Aldred et al. (2017)	Greater female preference for facilities with high level of separation from vehicular traffic	Significant
Jensen et al. (2019)	Female cyclists aversion to cycling in the wrong direction	Significant
Akar et al. (2013); Abasahl et al., 2018; Prati et al. (2019)	Greater female concern for safety due to vehicular traffic and a lack of bicycle lanes, paths, and trails, mainly for beginner cyclists	Significant
Grudgings et al. (2021)	Gender and age difference on cycling behaviors in cities without a well-developed network of bikeways	Significant
AitBihiOuali and Klingen (2022)	Dedicated cycling infrastructure that increases participation among female cyclists	Significant
Twaddle et al. (2010)	No significant gender difference in the selection of any of the cycling facilities	No significant
Carroll et al., 2020	No statistically significant gender differences in traffic risk aversion	No significant

more experienced and gender-specific preferences for bikeways gradually vanish.

With this perspective in mind, the present study pursues a more fine-grained approach, and the goal is to find answer to these research questions: i) what are the female cyclist's route attributes preferences with respect to male cyclists in a city with an already well-developed network of bikeways? and ii) what are the differences in route attribute preferences between young and old cyclists, as well as those between infrequent and frequent female cyclists?

It is worth clarifying that this study does not develop route choice models but explores potential differences between male and female cyclists; in particular, it was designed to explain the differences between gender-specific route attributes that male and female cyclists experience under similar conditions. The employed method for the present study is the Oaxaca-Blinder decomposition, which has been often adopted to analyze the effect of a specific discrimination factor (gender and race, among others) over an outcome of interest. The majority of applications of this technique may be found in the labor market and discrimination literature, where the primary differential outcome to explain is the wage gap (for meta-analyses, see Stanley and Jarrell, 1998, or Weichselbaumer and Winter-Ebmer, 2005).

3. Data set description and characterization

The present analysis is based on a large set of recorded GPS traces obtained from a data-collection campaign in the city of Bologna, called Bella Mossa (Città Metropolitana di Bologna, 2017). The recording of GPS traces during a "sustainable journey" (by bike, public transport, or walking) was performed on a volunteer basis but was incentivized by discounts at certain retailers. The campaign was conducted from April to September 2017, and resulted in approximately 270,000 recorded GPS traces from bike trips.[1] The study is focused only on trips recorded during weekday mornings (7 a.m.–10 a.m.), which are more likely linked to work trips, avoiding the noise generated by leisure trips, and also focusing on the more critical time of the day from a transportation perspective when traffic congestion is more likely to occur. Therefore, the total number of considered GPS traces has been 29,431, approximately 11% of all recorded trips. The Bella Mossa database contains information on various aspects of bicycle routes (e.g., start and end time of the trip, distance travelled, duration, and average speed) and some personal information regarding the recording cyclist (such as age, gender, and frequency of bike usage). The raw GPS traces consisted of geolocated points associated with timestamps; these points are usually recorded with an average frequency of a few seconds depending on the battery constraint of the used smartphone, whereas the app did not record data when it did not detect motion.

The data preprocessing step, based on the open-source software SUMOPy, is as follows (Rupi et al., 2020; Poliziani et al., 2022): the road network is imported from OpenStreetMap and manually corrected. GPS traces were successively imported and an initial cleaning filter eliminated invalid traces (those with either too few GPS points, too short distance or duration, or too high average velocity or with GPS points registered out of the study area; for thresholds, see Rupi et al., 2020, and Poliziani et al., 2022). Moreover, a specific point filter tries to estimate the initial and final position of the cyclist, facing two main problems: i) the cyclists activate the trip recording too early, resulting in a cloud of data points collected while the cyclist is moving around the bike for preliminary set up before the departure, and ii) if the GPS signal had been just turned on, it could require several seconds to estimate the correct smartphone position, therefore making the first point recorded far from all other points. The first problem was solved by deleting all points within 10 m of the first point. This latter is resolved by excluding the first point if it is further than 10 m from the second one. Some manual controls have proven that this threshold is more efficient.

Next, GPS traces were matched to the road network and analyzed together with the shortest routes connecting the same origins and destinations to quantify the variables used for the analysis. The formulation of the used algorithm has been presented by Schweizer et al. (2016), as well as how the considered methods and parameters affect the matched route; two control variables have also been provided to detect the quality of the matched routes, which allowed greater filtering to keep only the well-matched traces. In particular, we kept only routes with a ratio between the matched length and the direct length connecting all the GPS points between 80% and 120%; moreover, we kept only matched routes that were, on average, within 10 m with respect to the GPS points, considering the shortest distance between these latter and the matched links.

After all the filters were considered, only 16,300 GPS traces carried out by 2395 cyclists remained.

For the scope of this study, the selected dataset is split into a dataset for male cyclists (group $j = 0$) and a dataset for female cyclists (group $j = 1$). The male dataset contains $N_0 = 8859$ traces and the female dataset contains $N_1 = 7441$ traces (see Table 2).

The trip database is representative of the cyclist population (Poliziani et al., 2020), and women, despite being 53.8% of cyclists, travel 45.7% of the trips, as reported in Table 2. Female cyclists are observed to cycle on average less than males, but the number of trips by each cyclist is highly variable. A deeper insight highlights that gender groups of cyclists exhibit a different composition by age. Female cyclists represent 55% of the youngest cyclists (less than or equal to 35 years) and 51% of the oldest ones. Furthermore, data indicates that the number of trips is unevenly distributed across individuals: of cyclists who registered only a single trip during the entire campaign, 32.4% were female and 27.4% were male, whereas only 17.3% of female cyclists and 24.5% of male ones registered 10 or more trips; details are provided in Table A1. As a whole, the composition of gender groups is heterogeneous in many aspects and has to be analyzed in order to identify whether gender impacts route choices.

The list of personal characteristics and route attributes together with mean values and deviation by gender is shown in Table 3. Three categories have been considered: personal characteristics, origin/destination attributes, and gender-specific route attributes. *Personal characteristics* refer to individual factors (e.g., age, route length, and speed). *Origin/destination attributes* refer to factors characterizing the chosen route, are independent of gender and possible differences in the average values, and are due to trip origin/destination pairs. *Gender-specific route attributes* refer to factors characterizing the chosen route and are intended to show differences dependent on cyclist gender. The gender-specific differences are not evident simply by looking at the average attributes of the chosen route but rather at the averages of the attribute difference between the chosen and the respectively shortest route, as detailed further below.

The data analysis shows that female cyclists' routes cross the central

Table 2
Basic statistics: male and female dataset.

	Male ($j = 0$)	Female ($j = 1$)
Number of trips, N_j	8859	7441
Share of trips (%)	54.3	45.7
Number of cyclists	1107	1288
Share of cyclists (%)	46.2	53.8
Total number of trips per person	8.0	5.8
Total number of trips per person (SD)	11.5	7.9
Share of cyclists aged ≤35 years (%)	45.0	55.0
Share of cyclists aged >35 years (%)	49.0	51.0

[1] A trip is a movement from an origin to a destination in a given reference time by a given transportation mode, and it can have multiple alternative routes (or paths). In some cases (e.g., if the origin/destination pair and the time period are fixed and only a route is possible), "trip" and "route" are used interchangeably.

Table 3
Mean values and deviation of personal characteristics and route attributes.

Description	Symbol	Male ($j = 0$)		Female ($j = 1$)	
Personal characteristics		Mean $\overline{x}_m^0$	SD	Mean $\overline{x}_m^1$	SD
Age	A	35.2	12.2	34.1	11.7
Total number of trips per person	N_T	8.0	11.5	5.8	7.9
Average speed (m/s]	V_{AV}	3.9	1.2	3.3	0.9
Ratio between length of chosen route and shortest route (%)	R_{EXTRA}	122.6	70.0	121.9	19.9
Origin/destination attributes					
Route length (km)	L_T	2.9	1.8	2.8	1.6
Share of route length in city center (%)	S_{CENTER}	38.7	37.0	42.3	36.8
Share of mixed access bikeways[a] (%)	S_{MIX}	30.6	21.1	31.0	20.9
Share on low priority roads[b] (%)	$S_{LOWPRIO}$	69.8	25.8	70.0	25.7
Gender-specific route attributes		Mean $\overline{y}_k^0$	SD	Mean $\overline{y}_k^1$	SD
Intersection per km (1/km]	ρ_{NOD}	16.462	3.92	16.632	3.40
Average number of possible maneuvers per intersection[c]	N_{MAN}	10.392	1.53	10.384	1.34
Number of left turns at intersections per km (1/km)	ρ_{LEFT}	2.188	1.26	2.175	1.19
Number of right turns at intersections per km (1/km)	ρ_{RIGHT}	2.447	1.34	2.404	1.28
Number of straight crossings per km (1/km)	$\rho_{STRAIGHT}$	11.148	3.42	11.357	3.20
Number of turns made at intersections per km (1/km)	ρ_{TURN}	4.635	2.30	4.579	2.17
Number of intersections without traffic lights per km (1/km)	ρ_{NOTL}	13.581	3.73	13.805	3.42
Number of intersections with traffic lights per km (1/km)	ρ_{TL}	2.881	1.98	2.828	1.83
Number of link priority changes per km[d] (1/km)	ρ_{PRIOCH}	0.879	1.09	0.858	1.08

[a] Network links with reserved access for bikes and either pedestrians or bus.
[b] Share of kilometers on road links with priority less than 7 on a scale from 1 to 13, based on number of lanes and maximum allowed speed.
[c] Average number of maneuvers a cyclist can perform from any incoming link at the intersection.
[d] Number of times the road link priority (see above) changes per kilometer along the chosen route.

area of the city with a higher percentage than those of male cyclists (about 6% more); this could be related to the fact that center roads are generally more crowded and more illuminated, also for a high density of shops. However, this same aspect is not related to a high density of schools in the central area of the city; in this case, the percentage of cyclists (both genders) who accompany their children to school is very low (around 1%), as revealed by our monitoring activities, and the majority of cyclists bringing their children to school are male.

Table 3 also shows that women cycle at lower average speed (approximately 15% less) than male cyclists, as highlighted in other works (e.g., Rupi et al., 2019). The route length of the female cyclists is on average as long as the ones of male cyclists, with the same percentage of mixed- and low-priority roads. Moreover, the routes of both male and female cyclists are approximately 20% longer compared to the length of the shortest path, but the extra lengthening compared to the shortest path is characterized by a much higher standard deviation for men. Note that the mean values of gender-specific route attributes seem to show no remarkable differences. Fig. 1 depicts the density distribution of chosen (solid) and shortest (dash) routes for males and females: the more the chosen routes density is displayed to the right of the shortest routes density, the more is the routes are lengthened by the respective gender.

For the aims of the paper, it is necessary to clarify i) whether the differences in average gender-specific route attribute values stem from differences in the two datasets (e.g., if the majority of male cyclists lived in an area with higher intersection density and the majority of female cyclists lived in an area with low intersection density) or ii) whether male and female cyclists, moving between the same OD pair, deliberately chose a route with different attributes. The latter is often referred to as the *gender discrimination*.

In order to separate these two effects, the Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973) was employed (see Section 4).

4. Methodology

As previously introduced, the Oaxaca-Blinder decomposition is used to decompose the difference in an average outcome between male and female cyclists, which is also called the gender differential. The main idea is to model gender differential by analyzing two linear regressions between the male and female outcomes as dependent variables, (e.g., the gender-specific attributes) and the independent variables (e.g., the personal characteristics and origin/destination attributes). Then, the first regression is done with data recorded by male cyclists, and the second regression is done with data recorded by female cyclists. These two regressions result in a coefficient vector for male and female cyclists, respectively. The gender difference can be explained either by different values of the independent variables or by differences between the coefficient vector of the two linear regressions. The latter, which is defined as the gender gap, can be considered as an indicator that highlights the systematic differences between male and female cyclists, which is independent of the composition of the dataset.

For this analysis, the linear model between each of the $k = 1, \ldots, K$ outcomes $y^j_{k,i}$ of gender j and trip i and $m = 1, \ldots, M$ independent variables $x^j_{m,i}$ can be expressed by:

$$y^j_{ki} = \beta^j_{k0} + \sum_{m=1}^{M} x^j_{mi}\beta^j_{km} + \varepsilon^j_{ki}, E\left[\varepsilon^j_{ki}\right] = 0 \quad j \in (1,0) \tag{1}$$

where β^j_{km} is the coefficient of the mth independent variable to estimate the kth outcome and ε^j_{ki} is a stochastic, zero-mean error-component to account for differences between the real value y^j_{ki} and the linear estimation. Linear regressions are made for both gender j and independently for each of the K attributes by using the ordinary least squares estimator. The results are the estimated coefficients $\widehat{\beta}^j_{km}$ for gender j, outcome k, and independent variable m.

Assuming zero average error, the kth average estimated outcome, $\overline{y}^j_k$, can be expressed as:

$$\overline{y}^j_k = \frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{m=0}^{M} x^j_{mi}\widehat{\beta}^j_{km} = \overline{\boldsymbol{X}}^{j'}\widehat{\boldsymbol{\beta}}^j_k \tag{2}$$

where $\overline{\boldsymbol{X}}^j = [1, \overline{x}^j_1 \ldots \overline{x}^j_M]'$ is the vector containing M values of the average independent variables from gender j, whereas $\widehat{\boldsymbol{\beta}}^j_k = [\widehat{\beta}^j_{k0}, \widehat{\beta}^j_{k1} \ldots \widehat{\beta}^j_{kM}]'$ is the coefficient vector to estimate the k th outcome for gender j.

Then, the total mean gender differential of the k th gender specific attribute can be calculated by:

$$D_k = \overline{y}^1_k - \overline{y}^0_k = \left(\overline{\boldsymbol{X}}^1 - \overline{\boldsymbol{X}}^0\right)'\widehat{\boldsymbol{\beta}}^1_k + \overline{\boldsymbol{X}}^{0'}\left(\widehat{\boldsymbol{\beta}}^1_k - \widehat{\boldsymbol{\beta}}^0_k\right) \tag{3}$$

where the first term explains the effect of different compositions of the independent variables between groups and thus depends on their distribution over gender datasets. The second term, which is usually referred to the discrimination effect, explains the differences between groups associated with the same values of independent variables, e.g., $\overline{\boldsymbol{X}}^0$. The mean gender differential can also be expressed from the viewpoint of the alternative group $j = 0$, which in this case is:

$$D_k = \overline{y}^1_k - \overline{y}^0_k = \left(\overline{\boldsymbol{X}}^1 - \overline{\boldsymbol{X}}^0\right)'\widehat{\boldsymbol{\beta}}^0_k + \overline{\boldsymbol{X}}^{1'}\left(\widehat{\boldsymbol{\beta}}^1_k - \widehat{\boldsymbol{\beta}}^0_k\right) \tag{4}$$

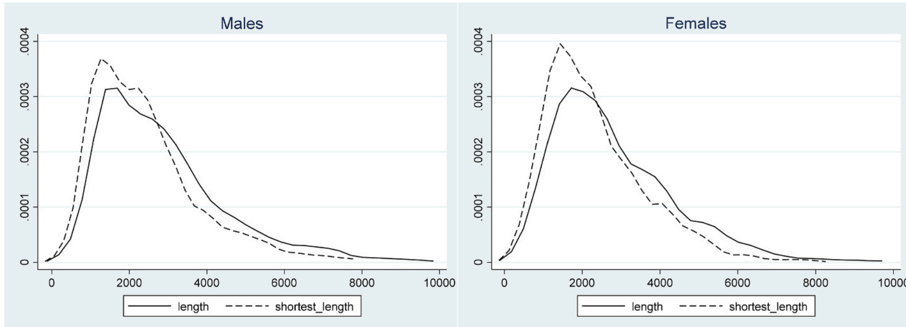

Fig. 1. Chosen vs shortest path length densities by gender.

Decompositions (3) and (4) are suitable for situations in which discrimination is directed only against one group. In the absence of specific reasons to assume discriminating coefficients, the following neutral specification[2] is generally adopted:

$$D_k = \bar{y}_k^1 - \bar{y}_k^0 = \left(\overline{X}^1 - \overline{X}^0\right)'\widehat{\beta}_k^* + \left(\overline{X}^{1'}\left(\widehat{\beta}_k^1 - \widehat{\beta}_k^*\right) + \overline{X}^{0'}\left(\widehat{\beta}_k^* - \widehat{\beta}_k^0\right)\right) \tag{5}$$

where $\widehat{\beta}_k^* = \alpha\,\widehat{\beta}_k^0 + (1-\alpha)\widehat{\beta}_k^1$ with $\alpha = 0.5$. In this symmetrical case, the gender gap related to the kth outcome is defined as:

$$G_k = \overline{X}^{1'}\left(\widehat{\beta}_k^1 - \widehat{\beta}_k^*\right) + \overline{X}^{0'}\left(\widehat{\beta}_k^* - \widehat{\beta}_k^0\right) \tag{6}$$

The Oaxaca-Blinder method allows for the control for the influence of the independent variables on the differences in gender specific attributes that are specified as outcomes. For instance, it may happen that in a dataset, the share of route length in city center (%) (S_{CENTER}) is over- or under-represented within one of the two gender groups. However, as long as the model contains S_{CENTER} as the independent variable, the Oaxaca-Blinder method will account for these differences; after controlling the mean gender differential for this independent variable, the mean residual differential part is determined by choices made by cyclists with respect to specific route attributes. Therefore, the gender gap, if present, is intended as the result of systematic choices of routes with certain attributes (e.g., lower intersection density) made by female cyclists with respect to male cyclists. This discrimination might be due to different levels of awareness and self-confidence and may restrict the choices from which female cyclists can choose their routes.

The study is designed to identify the differences in route attributes of the chosen routes (including the differences, gains, and losses with respect to the shortest route) while keeping personal characteristics and origin/destination attributes fixed. It is worthwhile to look at route choices made for the same location of trip OD to make them more comparable, although the probability of finding two trips having the same O and D for a given traveler is quite low even in large datasets. Although the ODs may be considered exogenous in relation to individual cyclist needs, the specific route chosen as conditional on the OD better expresses how cyclists choose routes and experience the related set of attributes, insofar as differences on outcomes related to the difference between chosen and shortest routes should isolate gender gap better than differences on outcomes related to the real paths, as these last may be affected by the different composition of ODs between female and male cyclists. Furthermore, some ODs could be characterized by the presence of few alternative itineraries, thus canceling the differences in the choice of the route due to gender.

5. Results of the Oaxaca-Blinder decomposition and policy implications

For this study, the nine gender-specific route attributes illustrated in Table 3 are analyzed as outcomes. They form a first set of outcomes. Moreover, to detect difference between genders, each gender-specific route attribute of the chosen route $y_{k,i}^j$, where $y_{k,i}^j$ is the k^{th} outcome from the i^{th} trip of gender j, is compared to the respective route attributes of the shortest route $w_{k,i}^j$, as suggested by Rupi et al. (2019). The outcome difference between chosen and shortest route attributes is defined as $\delta_{k,i}^j = y_{k,i}^j - w_{k,i}^j$, which leads to the following further three groups of outcome attributes in addition to the ones reported in Table 3:

$$y_{h+9,i}^j = \delta_{h,i}^j h = 1, \ldots\ldots, 9 \tag{7}$$

$$y_{h+18,i}^j = \begin{cases} -\delta_{h,i}^j \text{ if } & \delta_{h,i}^j < 0 \\ 0 \quad \text{if } & \delta_{h,i}^j \geq 0 \end{cases} h = 1, \ldots\ldots, 9 \tag{8}$$

$$y_{h+27,i}^j = \begin{cases} \delta_{h,i}^j \text{ if } & \delta_{h,i}^j \geq 0 \\ 0 \quad \text{if } & \delta_{h,i}^j < 0 \end{cases} .h = 1, \ldots\ldots, 9 \tag{9}$$

Outcomes (8) and (9) respectively determine the gains and the losses of attribute values with regard to the ones encountered on the shortest route.

In summary, four groups of outcomes that are gender- and route-specific were analyzed, and each group consisted of nine attributes, resulting in a total of 36 different outcomes.

Personal characteristics and the OD attributes are specified as independent variables. They are age, total number of trips per person, average speed (m/s), and route length (km), share of route length in city center (%), share of mixed access bikeways (%), and share on low-priority roads (%). Because it is expected that outcomes may display nonlinear responses to independent variables, their square transformations are also included. All the independent variables (linear and squared) are indicated as $x_{m,i}^j$, where mth indicates the independent variable revealed from trip i of gender j and $\bar{x}_m^j = \frac{1}{N_j}\sum_{i=1}^{N_j} x_{m,i}^j$ are the mean values averaged over the datasets of N_j trips.

The results, summarized in Table 4, are organized into two columns: the observed gender differential, D_k, and the gender gap, G_k, computed after having taken into account for the differences between groups of covariates composition (equation (3) and (6)). For both indicators, the mean value, standard deviation, p-value, and size effect (computed as

[2] Results are expected to be, to some extent, sensitive to the choice of coefficients and/or α. The choice of $\alpha = 0.5$, proposed by Reimers (1983), corresponds to working under the most neutral setting for the results. Of course, additional similar proposals for the choice of coefficients are available: Cotton (1988) suggests weighting the coefficients by the group sizes n0 and n1, while Neumark (1988) advocates the usage of the coefficients $\widehat{\beta}_k^*$ estimated from a pooled regression over both groups. While not all αs are interpretable, it is of interest to evaluate a subset of interesting cases. For this reason, in Table A4 we present results provided by a set of choices for the parameters' vector.

Table 4
Decompositions over original dataset: female versus male cyclists.

k	Route attribute	Exp. sign	Gender differential, ***D_k***	SE	p-Value	Eval	% Effect	Gender gap, ***G_k***	SE	p-Value	Eval	% Effect
1	ρ_{NOD}	-	0.170	0.057	0.001		1.0	0.103	0.055	0.029		0.6
2	N_{MAN}	-	−0.008	0.022	0.363		−0.1	0.024	0.022	0.135		0.2
3	ρ_{LEFT}	-	−0.013	0.019	0.253		−0.6	−0.045	0.017	0.005		−2.1
4	ρ_{RIGHT}	−	−0.043	0.021	0.018		−1.8	−0.072	0.019	0.000	es*	−3.0
5	$\rho_{STRAIGHT}$	+	0.209	0.052	0.000	es*	1.9	0.218	0.044	0.000	es*	1.9
6	ρ_{TURN}	-	−0.056	0.035	0.054		−1.2	−0.117	0.031	0.000	es*	−2.5
7	ρ_{NOTL}	-	0.224	0.056	0.000	us	1.6	0.199	0.055	0.000	us	1.5
8	ρ_{TL}	+	−0.053	0.030	0.037		−1.9	−0.095	0.023	0.000	us	−3.3
9	ρ_{PRIOCH}	-	−0.021	0.017	0.113		−2.4	−0.038	0.016	0.009		−4.3
	NG/NG*					1/2					3/5	
k	**Real to shortest**	**Exp. sign**	**Gender differential, *D_k***	**SE**	**p-Value**	**Eval**	**%Effect**	**Gender gap, *G_k***	**SE**	**p-Value**	**Eval**	**%Effect**
1	ρ_{NOD}	-	−0.162	0.051	0.001	es*	−36.6	−0.242	0.045	0.000	es*	−54.7
2	N_{MAN}	-	0.007	0.017	0.330		3.0	0.029	0.017	0.043		11.6
3	ρ_{LEFT}	-	−0.077	0.017	0.000	es*	−14.5	−0.079	0.017	0.000	es*	−14.9
4	ρ_{RIGHT}	−	−0.040	0.019	0.016		−7.0	−0.056	0.018	0.001	es*	−9.7
5	$\rho_{STRAIGHT}$	+	−0.018	0.046	0.343		−1.5	−0.058	0.044	0.094		−4.6
6	ρ_{TURN}	-	−0.117	0.030	0.000	es*	−10.6	−0.134	0.030	0.000	es*	−12.2
7	ρ_{NOTL}	-	−0.257	0.046	0.000	es*	−40.7	−0.268	0.045	0.000	es*	−42.4
8	ρ_{TL}	+	0.095	0.031	0.001	us	50.0	0.026	0.019	0.084		13.7
9	ρ_{PRIOCH}	-	−0.101	0.022	0.000	es*	−1178.1	−0.116	0.022	0.000	es*	−1358.3
	NG/NG*					5/6					6/6	
k	**Gains**	**Exp. sign**	**Gender differential, *D_k***	**SE**	**p-Value**	**Eval**	**%Effect**	**Gender gap, *G_k***	**SE**	**p-Value**	**Eval**	**% Effect**
1	ρ_{NOD}	+	0.033	0.040	0.204		2.7	0.097	0.031	0.001	es*	7.8
2	N_{MAN}	+	0.009	0.009	0.156		4.0	0.007	0.009	0.235		2.8
3	ρ_{LEFT}	+	0.018	0.008	0.010		11.0	0.018	0.008	0.013		10.7
4	ρ_{RIGHT}	+	0.001	0.009	0.463		0.5	0.005	0.009	0.281		2.8
5	$\rho_{STRAIGHT}$	−	0.033	0.036	0.184		1.9	0.052	0.035	0.068		2.9
6	ρ_{TURN}	+	0.017	0.014	0.113		6.4	0.021	0.014	0.058		8.2
7	ρ_{NOTL}	+	0.093	0.033	0.003		6.8	0.117	0.033	0.000	es*	8.5
8	ρ_{TL}	−	−0.072	0.027	0.003		−23.2	−0.017	0.011	0.054		−5.5
9	ρ_{PRIOCH}	+	0.063	0.016	0.000	es*	16.0	0.073	0.017	0.000	es*	18.6
	NG/NG*					1/1					3/3	
k	**Losses**	**Expcted sign**	**Gender differential, *D_k***	**SE**	**p-Value**	**Eval**	**% Effect**	**Gender Gap, *G_k***	**s.e**	**p-Value**	**Eval**	**% effect**
1	ρ_{NOD}	-	−0.128	0.023	0.000	es*	−15.9	−0.144	0.023	0.000	es*	−17.8
2	N_{MAN}	-	0.017	0.012	0.081		3.5	0.035	0.012	0.001	us	7.3
3	ρ_{LEFT}	-	−0.058	0.013	0.000	es*	−8.4	−0.061	0.013	0.000	es*	−8.8
4	ρ_{RIGHT}	−	−0.039	0.014	0.002		−5.2	−0.050	0.014	0.000	es*	−6.7
5	$\rho_{STRAIGHT}$	+	0.014	0.018	0.215		2.8	−0.006	0.018	0.369		−1.2
6	ρ_{TURN}	-	−0.100	0.024	0.000	es*	−7.3	−0.113	0.023	0.000	es*	−8.3
7	ρ_{NOTL}	-	−0.164	0.023	0.000	es*	−22.2	−0.151	0.022	0.000	es*	−20.5
8	ρ_{TL}	+	0.023	0.014	0.045		4.6	0.009	0.013	0.247		1.8
9	ρ_{PRIOCH}	-	−0.038	0.011	0.000	es*	−9.4	−0.043	0.011	0.000	es*	−10.7
	NG/NG*					5/5					6/7	
Total NG/NG*					12/14					18/21		

NOTE: Exp. sign = Expected sign; u = sign unexpected; e = expected with respect to the hypothesis; s = significant based on p-values with the Bonferroni correction; * = expected and significant.

differential or gap, divided by the average outcome value $\overline{y}_k^j$) are also provided. Because the same model is estimated over the same data, the significance has been adjusted with the Bonferroni correction, and only empirical p-values lower than the adjusted theoretical p-value are considered significant (see discussion further below). Moreover, being in the presence of a large sample, the standard errors become extremely small, such that even very small differences between the estimate and the null hypothesis become statistically significant (Good, 1992; Tukey, 1991). No solutions exist for this issue and, following the suggestions of literature on the topic (Lin et al., 2013), together with statistical significance of results, we will provide evaluations on their practical significance by assessing the magnitude of size effects. For the $k^{\text{-th}}$ outcome, the size effect is computed as the ratio of the specific difference and gap to the average value of the outcome in percentage scale; hence, $\frac{D_k}{\overline{y}_k} \times 100$ and $\frac{G_k}{\overline{y}_k} \times 100$.

In Table 4, the outcomes are grouped into chosen route attributes, difference between chosen route and shortest route attribute, and gains and losses with respect to the shortest routes, as defined in Section 3. The first observation is that for the first group gender differential and gender gap are small with respect to the absolute outcome value. Only in the second group, where the route attributes are compared with the route attributes of the shortest path, are the magnitude of D_k and G_k markedly higher. The same is true for the two remaining outcome groups, gains and losses, with respect to the shortest route.

This raises the question of whether there is a pattern in D_k and G_k indicating the preferences of female cyclists. In other words, what are the differences in reasons why female and male cyclists make deviations with respect to the shortest route? Such a pattern could be identified by introducing the concept of route complexity, which has been implicitly determined as a result of the interaction of several elements; in particular, the route complexity increases for an increasing average number of possible maneuvers per intersection, number of intersections per kilometer, number of left and right turns at intersections per kilometer, total number of turns made at intersections per kilometer, number of intersections without traffic lights per kilometer, and number of link priority changes per kilometer. However, the route complexity decreases with an increasing number of straight crossings per kilometer and with an increasing number of intersections with traffic lights per kilometer. Straight crossing maneuvers, particularly at traffic light intersections, could be perceived as if there is no real intersection.

It is worthwhile to note that, in this preliminary phase, route complexity has not been mathematically formulated; in the following, we refer to the interaction of the above variables in order to deduce a qualitative level of complexity.

The pattern seen in D_k and G_k values could lead to the hypothesis that female cyclists seek less-complex routes with respect to their male counterparts. If this were the case, the D_k and G_k would be negative for all the route complexity increasing outcomes in the route attributes group simply because male cyclists would choose routes that were, on average, more complex. Also, for the outcomes of the chosen versus the shortest route, D_k and G_k are expected to be negative for all outcomes in which route complexity increases because per the hypothesis, female cyclists seek less-complex routes (or simpler routes) with respect to the shortest routes compared with male cyclists. Instead, in the gains group, D_k and G_k are expected to be positive because female cyclists seek a higher gain in less-complex routes with respect to male cyclists. In the losses outcome group, D_k and G_k are expected to show a negative sign for outcomes in which route complexity increases because female cyclists would seek lower losses in less-complex routes with respect to the shortest routes than male cyclists. The expected signs of D_k and G_k for route complexity decreasing outcomes is just the opposite of the sign of outcomes in which route complexity increases within each outcome group.

In order to confirm or reject the hypothesis that female cyclists seek less-complex routes, the following evaluation is considered within each outcome group: *NG* is the number of significant outcomes in a sense that the outcome respects the Bonferroni-corrected p-values (see "Eval" columns in Table 4) and *NG** is the number of significant outputs with the expected sign of D_k or G_k.

From Table 4, *NG, NG*,* and the size effects are considerably higher for the outcome groups where the attributes are compared with the shortest route, which justifies this differential analysis. Note that most of the significant outcomes are also expected, e.g., *NG* is equal to or greater by one than *NG**. It is further obvious that the gender gap G_k delivers more significant results with respect to the gender differential D_k, which means the Oaxaca-Blinder decomposition introduces differences in evidence that would remain hidden otherwise. A considerable number of outcomes are significant and have a relevant size effect, confirming the hypothesis that female cyclists prefer simpler/less-complex routes. This aspect is surely important also in terms of policy-level implications because planners are asked to avoid infrastructure elements leading to route complexity, such as a high number of intersections, particularly uncontrolled ones, as well as a high number of turning maneuvers, which may prevent some part of potential travelers to use the bike.

The outcomes with the largest size effects (in the "chosen versus shortest" outcome group) are the density of priority changes, ρ_{PRIOCH} (<-100.0%); the intersections density, ρ_{NOD} (−54.7%); and the intersection density without traffic lights, ρ_{NOTL} (−42.4%). In practical terms, during the planning phase, it is important to plan cycling networks with as many more-continuous bike pathways (i.e., with a small number of conflicting points) as possible in order to guarantee the usability of the routes to all categories of users.

Consistently within expectations, the size effect of left turn intersection density, ρ_{LEFT} (−14.9%), is greater than the size effect of right-turn intersections, ρ_{RIGHT} (−9.7%). Left turns at intersections have a higher number of potential conflict points with respect to right turns and are therefore considered more complex and are more avoided by female cyclists. This aspect has an important planning implication because left turns are the riskiest manoeuvres and should be avoided as much as possible. In particular, protected traffic lights phases could be planned for cyclists, together with some infrastructure elements such as reserved areas in front of cars and speed bumpers for vehicles in case of high road traffic flows. However the density of crossings passed in a straight path gives insignificant results.

To gain more insight, gender gaps have been estimated in subsamples of the dataset derived by classes of age and of trip frequency (see Table 2). In particular, light users are those who recorded fewer than 8 trips during the entire campaign, whilst frequent users recorded more than 8 trips. Although frequent bicycle users may have gained a deeper knowledge of route network attributes, light cyclists might have not experienced all route alternatives.

Table 5 provides the following evidence for gender gaps G_k (gender differentials D_k are omitted because gender gaps are confirmed to be better observed after controlling for the data composition); there are more significant gender gaps for young cyclists with respect to older cyclists. Likewise, frequent users show more significant outcomes than light cyclists. These findings are briefly summarized again by using the previously defined ratio *NG*/NG*. In general, the gender gap has more significant and expected outcomes for younger users (8 of 8 for light cyclists and 23 of 23 for frequent cyclists), whereas elderly cyclists have less significant outcomes (0 of 4 for light cyclists and 2 of 10 for frequent cyclists). A higher propensity to cycling has been detected for younger rather than older women, which suggests different behaviors between young and older cyclists, as also shown by other studies focusing on behavioral analyses of young bike users (e.g., Picasso et al., 2020).

Both light young and frequent young female cyclists make deviations to make routes less complex: in both cases, the effects of gender gaps are of relevant size, up to a gap of 28.8% of gain in turns for young light female cyclists, and 37.8% of gain in priority changes for young frequent cyclists. In this last group, the size effects of gaps in losses are even greater than averaged for all female cyclists; female cyclists reduce intersection density (number of intersections per kilometer) by half compared to male cyclists in terms of total intersections (−57.5%) and intersections without traffic lights (−57.6%) and, in general, they avoid more than 20% of increased turns and crossings and one-third of the priority changes. Some of these results have already been highlighted in Lusk et al. (2014).

To further validate our results, we provide some robustness checks. Firstly, in light of the fact that of about 30% cyclists registered only one trip, we excluded those cyclists from our analysis. As a whole, results are coherent with the previous ones indicating that the possible difficulty in using the app did not act as a strong selection mechanism (see Tables A2 and A3). Secondly, we repeat the decompositions by choosing different parameters' vectors, corresponding to a number of a priori assumption about discrimination (Tables A4). Results change only negligibly, i.e., do not change neither under the less realistic hypothesis of discrimination only against male cyclists. In any other circumstance, the presented results hold true.

Furthermore, in order to evaluate the gender-specific effect in relation to route length and average speed, a similar analysis has been run, and two additional variables were computed: low (less than 3.5 m/s) and high (more than 3.5 m/s) average speed and shorter (less than 2.5 km) and longer (more than 2.5 km) routes. Thus, there emerged a higher gender gap when longer rather shorter routes are considered and, similarly, low speed rather than high speed (see Table A5).

All these findings support the hypothesis that the behavior of older female cyclists is more similar to that of male cyclists. This could be either because older female cyclists gained more experience and do not shy away from complex junctions or because of a kind of self-selection in which more-anxious female cyclists gave up cycling.

To summarize, also in terms of policy implications, to spread the usability of the routes to all categories of users, some key planning factors have emerged from the above analyses, particularly those reducing i) intersection density, ii) number of uncontrolled intersections, and iii) left turns.

6. Conclusions

This study was conducted to investigate the route attribute differences for male and female cyclists in a city with a mature network of bikeways based on real data. The large number of detailed route

Table 5
Decompositions by age and frequency: female versus male cyclists.

			Light cyclists						Frequent cyclists					
			Young			Older			Young			Older		
K	Route attribute	Exp. sign	Gender gap, ***G_k***	Eval	% Effect	Gende gap, ***G_k***	Eval	% Effect	Gender gap, ***G_k***	Eval	% Effect	Gender gap, ***G_k***	Eval	% Effect
1	ρ_{NOD}	-	−0.090		*−0.5*	0.611	us	*3.7*	−0.441	es*	*−2.7*	0.375	us	*2.3*
2	N_{MAN}	-	0.032		*0.3*	0.163	us	*1.6*	−0.072		*−0.7*	0.038		*0.4*
3	ρ_{LEFT}	-	−0.103	es*	*−4.6*	−0.030		*−1.4*	−0.111	es*	*−5.2*	0.043		*1.9*
4	ρ_{RIGHT}	−	−0.067		*−2.8*	−0.020		*−0.8*	−0.269	es*	*−11.4*	−0.010		*−0.4*
5	$\rho_{STRAIGHT}$	+	0.082		*0.7*	0.670	es*	*6.1*	−0.043		*−0.4*	0.325	us	*2.9*
6	ρ_{TURN}	-	−0.170	es*	*−3.6*	−0.051		*−1.1*	−0.380	es*	*−8.4*	0.033		*0.7*
7	ρ_{NOTL}	-	0.142		*1.0*	0.575	us	*4.2*	0.033		*0.2*	0.008		*0.1*
8	ρ_{TL}	+	−0.233	es*	*−7.5*	0.035		*1.4*	−0.474	es*	*−15.2*	0.368	es*	*14.4*
9	ρ_{PRIOCH}	-	0.006		*0.7*	0.031		*3.7*	−0.053		*−5.5*	−0.115	es*	*−14.1*
	NG/NG*			3/3			1/4			5/5			2/4	
k	Real to shortest	Exp. sign	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect
1	ρ_{NOD}	-	−0.327	es*	.	−0.032		*−5.4*	−0.717	es*	.	0.065		*12.1*
2	N_{MAN}	-	−0.026		*−13.1*	0.034		*12.9*	−0.009		*−3.6*	0.134	us	*45.9*
3	ρ_{LEFT}	-	−0.072		*−13.9*	0.005		*0.8*	−0.148	es*	*−29.8*	−0.095		*−16.5*
4	ρ_{RIGHT}	−	−0.063		*−10.8*	0.018		*3.1*	−0.213	es*	*−44.7*	−0.039		*−6.1*
5	$\rho_{STRAIGHT}$	+	−0.124		*−13.7*	0.108		*7.4*	−0.336	es*	*−27.9*	0.160		*10.6*
6	ρ_{TURN}	-	−0.135		*−12.2*	0.023		*2.0*	−0.361	es*	*−37.1*	−0.134		*−11.0*
7	ρ_{NOTL}	-	−0.269	es*	*−65.0*	−0.036		*−4.8*	−0.624	es*	*−69.2*	−0.235		*−44.1*
8	ρ_{TL}	+	−0.058		*−20.8*	0.004		*2.6*	−0.093		*−29.5*	0.300	es*	.
9	ρ_{PRIOCH}	-	−0.078		.	0.023		*30.1*	−0.312	es*	.	−0.119		.
	NG/NG*			2/2			0/0			7/7			1/2	
k	Gains	Exp. sign	**Gender Gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect
1	ρ_{NOD}	+	0.272	es*	*26.2*	−0.006		*−0.4*	0.303	es*	*23.1*	−0.177		*−13.1*
2	N_{MAN}	+	0.028		*12.2*	0.047		*19.5*	0.002		*1.0*	−0.043		*−16.7*
3	ρ_{LEFT}	+	0.040		*23.1*	0.004		*2.2*	−0.001		*−0.8*	0.017		*9.5*
4	ρ_{RIGHT}	+	0.024		*14.1*	−0.025		*−13.9*	0.028		*13.1*	0.008		*4.4*
5	$\rho_{STRAIGHT}$	−	0.125		*8.2*	−0.002		*−0.1*	0.227	es*	*13.5*	−0.186		*−9.5*
6	ρ_{TURN}	+	0.073	es*	*28.8*	−0.026		*−9.8*	0.022		*8.1*	0.018		*6.9*
7	ρ_{NOTL}	+	0.229	es*	*19.4*	0.028		*1.9*	0.277	es*	*18.4*	−0.008		*−0.6*
8	ρ_{TL}	−	−0.003		*−1.1*	0.061		*21.8*	0.012		*5.2*	−0.159	us	*−36.5*
9	ρ_{PRIOCH}	+	0.039		*11.5*	0.012		*3.9*	0.169	es*	*37.8*	0.128		*27.6*
	NG/NG*			3/3			0/0			4/4			0/1	
k	Losses	Exp. sign	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect
1	ρ_{NOD}	-	−0.055		*−6.1*	−0.038		*−4.9*	−0.414	es*	*−57.5*	−0.112		*−13.7*
2	N_{MAN}	-	0.002		*0.5*	0.081		*16.0*	−0.006		*−1.4*	0.092	us	*16.7*
3	ρ_{LEFT}	-	−0.032		*−4.7*	0.008		*1.2*	−0.149	es*	*−23.1*	−0.079		*−10.4*
4	ρ_{RIGHT}	−	−0.040		*−5.3*	−0.007		*−0.9*	−0.185	es*	*−27.0*	−0.031		*−3.8*
5	$\rho_{STRAIGHT}$	+	0.001		*0.2*	0.107		*22.1*	−0.109	es*	*−22.9*	−0.026		*−5.7*
6	ρ_{TURN}	-	−0.062		*−4.6*	−0.003		*−0.2*	−0.339	es*	*−27.1*	−0.116		*−7.9*
7	ρ_{NOTL}	-	−0.040		*−5.3*	−0.008		*−1.1*	−0.347	es*	*−57.6*	−0.243	es*	*−28.6*
8	ρ_{TL}	+	−0.061		*−10.9*	0.065		*15.5*	−0.080		*−14.5*	0.141	es*	*32.7*
9	ρ_{PRIOCH}	-	−0.039		*−10.1*	0.035		*8.9*	−0.143	es*	*−33.9*	0.009		*2.2*
	NG/NG*			0/0			0/0			7/7			2/3	
Total NG/NG*				8/8			1/4			23/23			5/10	

attributes combined with a large set of recorded GPS traces has enabled a fine-grained study of the route preferences of female cyclists. In addition to the attributes of the chosen routes, the difference between the chosen and the shortest routes were analyzed as additional attributes to investigate why male and female cyclists would make detours. The Oaxaca-Blinder decomposition was employed to investigate the cause of the average difference in route attribute, also known as the gender differential, which should be independent of the composition of the dataset in terms of personal characteristics and origin/destination attributes.

The goal of the study was to understand potential gender gaps and gender differentials that may lead to different route choices by male and female cyclists, thus helping policy-makers and designers in planning more-suitable bike networks.

A key finding is that female cyclists tend to avoid complex route elements more than men, depending on both infrastructure and operational characteristics, such as high density of intersections and intersections without traffic lights, change in road type, and left turns. In other words, the study shows that, on average, female cyclists seem to prefer routes with as little interference as possible from other traffic streams, as deduced by their reluctance to use roads with numerous intersections and unsafe maneuvers such as left turns. However, the behavior of female cyclists is not homogenous among age groups; differences in preferred route attributes are more pronounced with young cyclists and seem to almost vanish for older cyclists. This seems to suggest that more-experienced female cyclists behave more similarly to their male counterparts, and differences tend to be less relevant with increased level of confidence in and knowledge of cycling. Reducing route complexity appears to be an important factor, particularly among young female cyclists, who prefer extending their route to meet such needs. In terms of policy implication, this suggests improving the features of the bikeways in order to reduce factors such as average number of possible maneuvers per intersection, number of intersections per kilometer, number of left turns and right turns at intersections per kilometer, total number of turns made at intersections per kilometer, number of intersections without traffic lights per kilometer, and number of link priority changes per kilometer. It is worth noting that all of these factors might be associated with the concept of safety when multimode traffic streams cross each other. Ultimately, female cyclists, on average, seem to be more influenced by safety issues than their male counterparts.

This analysis has deep implications and can influence policies regarding rethinking both infrastructure and network operational bike features in order to increase route usability for all categories of users. However, it also introduces an issue regarding route complexity, which in this preliminary phase has not been explicitly formulated. Future research will refer to a quantitative definition of route complexity, generating evidence for the most relevant involved variables and their different contributions in defining it. Furthermore, being that the analyzed dataset is composed of only of trips during the peak morning hours, it would be interesting to examine route attribute preferences for all types of trips, including multimode bike-sharing opportunities depending on multimode terminal locations, which requires further knowledge and availability of additional attributes such as road illumination and terminal locations.

Author statement

Rupi Federico: Conceptualization, Methodology, Writing- Original draft preparation, Writing- Reviewing and Editing. Freo Marzia: Methodology, Investigation, Data curation, Writing- Original draft preparation, Reviewing and Editing. Poliziani Cristian: Visualization, Investigation, Data curation, Writing- Original draft preparation, Reviewing and Editing. Postorino Maria Nadia: Methodology, Supervision, Writing- Original draft preparation, Writing- Reviewing and Editing. Schweizer Joerg: Methodology, Supervision, Data curation, Validation, Writing- Original draft preparation, Writing- Reviewing and Editing.

Data availability

The authors do not have permission to share data.

Appendix

Table A1
Distribution of cyclists by number of trips

Trips	Cyclists		Percent	
	Female	Male	Female	Male
1	417	303	32.4	27.4
2	197	140	15.3	12.6
3	123	95	9.5	8.6
4	86	84	6.7	7.6
5	61	64	4.7	5.8
6	63	43	4.9	3.9
7	49	50	3.8	4.5
8	40	27	3.1	2.4
9	29	30	2.3	2.7
10	27	28	2.1	2.5
>10	196	243	15.2	22.0
Total	1288	1107	100.0	100.0

Table A2
Decompositions over original dataset: female versus male cyclists (registering more than one trip)

k	Route attribute	Exp. sign	Gender differential, ***D_k***	SE	p-value	Eval	Gender gap, ***G_k***	SE	p-value	Eval
1	ρ_{NOD}	-	0.167	0.059	0.004		0.103	0.056	0.066	
2	N_{MAN}	-	−0.010	0.023	0.654		0.022	0.022	0.336	
3	ρ_{LEFT}	-	−0.010	0.020	0.608		−0.042	0.018	0.018	
4	ρ_{RIGHT}	–	−0.046	0.021	0.030		−0.073	0.019	0.000	es*
5	$\rho_{STRAIGHT}$	+	0.207	0.053	0.000	es*	0.216	0.045	0.000	es*
6	ρ_{TURN}	-	−0.056	0.036	0.120		−0.115	0.032	0.000	es*
7	ρ_{NOTL}	-	0.223	0.057	0.000	us	0.201	0.056	0.000	us*
8	ρ_{TL}	+	−0.056	0.031	0.068		−0.098	0.024	0.000	us*
9	ρ_{PRIOCH}	-	−0.022	0.017	0.206		−0.040	0.016	0.014	
	NG/NG*					1/2				3/5
k	Real to shortest	Exp. sign	**Gender differential, *D_k***	SE	p-value	Eval	**Gender gap, *G_k***	SE	p-value	Eval
1	ρ_{NOD}	-	−0.178	0.053	0.001	es*	−0.255	0.046	0.000	es*
2	N_{MAN}	-	0.005	0.017	0.776		0.028	0.017	0.106	
3	ρ_{LEFT}	-	−0.081	0.018	0.000	es*	−0.084	0.017	0.000	es*
4	ρ_{RIGHT}	–	−0.040	0.019	0.036		−0.055	0.019	0.003	
5	$\rho_{STRAIGHT}$	+	−0.041	0.047	0.382		−0.077	0.045	0.084	
6	ρ_{TURN}	-	−0.121	0.031	0.000	es*	−0.139	0.030	0.000	es*
7	ρ_{NOTL}	-	−0.278	0.047	0.000	es*	−0.283	0.046	0.000	es*
8	ρ_{TL}	+	0.100	0.032	0.002		0.028	0.019	0.144	
9	ρ_{PRIOCH}	-	−0.108	0.022	0.000	es*	−0.122	0.023	0.000	es*

(*continued on next page*)

Table A2 (*continued*)

k	Real to shortest	Exp. sign	**Gender differential, D_k**	SE	p-value	Eval	**Gender gap, G_k**	SE	p-value	Eval
	NG/NG*					5/5				5/5
k	**Gains**	**Exp. sign**	**Gender differential, D_k**	**SE**	**p-Value**	**Eval**	**Gender gap, G_k**	**SE**	**p-Value**	**Eval**
1	ρ_{NOD}	+	0.042	0.041	0.313		0.104	0.032	0.001	es*
2	N_{MAN}	+	0.012	0.009	0.206		0.009	0.009	0.363	
3	ρ_{LEFT}	+	0.021	0.008	0.009		0.021	0.008	0.010	
4	ρ_{RIGHT}	+	0.002	0.010	0.801		0.007	0.009	0.479	
5	$\rho_{STRAIGHT}$	–	0.051	0.037	0.168		0.068	0.035	0.053	
6	ρ_{TURN}	+	0.020	0.014	0.157		0.025	0.014	0.075	
7	ρ_{NOTL}	+	0.104	0.034	0.002		0.125	0.034	0.000	es*
8	ρ_{TL}	–	−0.077	0.027	0.005		−0.019	0.011	0.071	
9	ρ_{PRIOCH}	+	0.069	0.017	0.000	es*	0.077	0.018	0.000	es*
	NG/NG*					1/1				3/3
k	**Losses**	**Exp. sign**	**Gender differential, D_k**	**SE**	**p-Value**	**Eval**	**Gender gap, G_k**	**SE**	**p-Value**	**Eval**
1	ρ_{NOD}	-	−0.136	0.024	0.000	es*	−0.151	0.023	0.000	es*
2	N_{MAN}	-	0.017	0.012	0.165		0.036	0.012	0.002	
3	ρ_{LEFT}	-	−0.060	0.014	0.000	es*	−0.063	0.013	0.000	es*
4	ρ_{RIGHT}	–	−0.037	0.014	0.008		−0.049	0.014	0.001	es*
5	$\rho_{STRAIGHT}$	+	0.010	0.018	0.572		−0.009	0.018	0.624	
6	ρ_{TURN}	-	−0.101	0.024	0.000	es*	−0.114	0.024	0.000	es*
7	ρ_{NOTL}	-	−0.174	0.023	0.000	es*	−0.158	0.022	0.000	es*
8	ρ_{TL}	+	0.023	0.014	0.096		0.009	0.014	0.523	
9	ρ_{PRIOCH}	-	−0.039	0.011	0.000	es*	−0.045	0.011	0.000	es*
	NG/NG*					5/5				6/6
Total NG/NG*					12/14				17/19	

NOTE: Exp. sign = Expected sign; u = sign unexpected; e = expected with respect to the hypothesis; s = significant based on p-values with the Bonferroni correction; * = expected and significant.

Table A3
Decompositions by class of age and frequency: female versus male cyclists (registering more than one trip)

			Light cyclists			
k	Route attribute	Exp. sign	Young Gender gap, G_k	Eval	Older Gender gap, G_k	Eval
1	ρ_{NOD}	-	−0.088		0.614	us
2	N_{MAN}	-	0.030		0.166	
3	ρ_{LEFT}	-	−0.105	es*	−0.018	
4	ρ_{RIGHT}	–	−0.068		−0.019	
5	$\rho_{STRAIGHT}$	+	0.087		0.663	es*
6	ρ_{TURN}	-	−0.173		−0.038	
7	ρ_{NOTL}	-	0.158		0.574	es*
8	ρ_{TL}	+	−0.246	us	0.040	
9	ρ_{PRIOCH}	-	0.003		0.035	
	NG/NG*			1/2		2/3
k	Real to shortest	Exp. sign	**Young Gender gap, G_k**	Eval	**Older Gender gap, G_k**	Eval
1	ρ_{NOD}	-	−0.344	es*	−0.084	
2	N_{MAN}	-	−0.034		0.038	
3	ρ_{LEFT}	-	−0.088		0.012	
4	ρ_{RIGHT}	–	−0.073		0.036	
5	$\rho_{STRAIGHT}$	+	−0.141		0.024	
6	ρ_{TURN}	-	−0.161		0.048	
7	ρ_{NOTL}	-	−0.292	es*	−0.076	
8	ρ_{TL}	+	−0.053		−0.008	
9	ρ_{PRIOCH}	-	−0.094		0.036	
	NG/NG*			2/2		0/0
k	Gains	Exp. sign	**Young Gender gap, G_k**	Eval	**Older Gender gap, G_k**	Eval
1	ρ_{NOD}	+	0.292	es*	0.031	
2	N_{MAN}	+	0.038		0.050	
3	ρ_{LEFT}	+	0.048	es*	0.010	
4	ρ_{RIGHT}	+	0.031		−0.028	
5	$\rho_{STRAIGHT}$	–	0.151		0.066	
6	ρ_{TURN}	+	0.087	es*	−0.022	
7	ρ_{NOTL}	+	0.247	es*	0.068	
8	ρ_{TL}	–	−0.010		0.069	
9	ρ_{PRIOCH}	+	0.050		0.006	
	NG/NG*			4/4		0/0
k	Losses	Exp.sign	**Young Gender gap, G_k**	Eval	**Older Gender gap, G_k**	Eval
1	ρ_{NOD}	-	−0.052		−0.053	
2	N_{MAN}	-	0.004		0.087	
3	ρ_{LEFT}	-	−0.039		0.021	
4	ρ_{RIGHT}	–	−0.042		0.008	

(*continued on next page*)

Table A3 (*continued*)

k	Losses	Exp.sign	**Young Gender gap, G_k**	Eval	**Older Gender gap, G_k**	Eval
5	$\rho_{STRAIGHT}$	+	0.010		0.090	
6	ρ_{TURN}	-	−0.074		0.026	
7	ρ_{NOTL}	-	−0.045		−0.008	
8	ρ_{TL}	+	−0.063		0.061	
9	ρ_{PRIOCH}	-	−0.043		0.042	
	NG/NG*			0/0		0/0
	Total NG/NG*		7/8		2/3	

Table A4
Decompositions using different coefficients' vectors: female versus male cyclists

Specifications*		Against female cyclists		Slightly against female cyclists		No discrimination		Against male cyclists	
		(1)		(2)		(3)		(4)	
Route attributes	Exp. sign	Gap G_k		Gap G_k		Gap G_k		Gap G_k	
ρ_{NOD}	−	0.097		0.093		0.103		0.045	
N_{MAN}	−	0.031		0.022		0.024		−0.024	
ρ_{LEFT}	−	−0.050		−0.046		−0.045		−0.077	es*
ρ_{RIGHT}	−	−0.071	es*	−0.072	es*	−0.072	es*	−0.114	es*
$\rho_{STRAIGHT}$	+	0.217	es*	0.210	es*	0.218	es*	0.236	es*
ρ_{TURN}	−	−0.121	es*	−0.118	es*	−0.117	es*	−0.191	es*
ρ_{NOTL}	−	0.206	us	0.184	us	0.199	us	0.102	
ρ_{TL}	+	−0.108	us	−0.091	us	−0.095	us	−0.056	
ρ_{PRIOCH}	−	−0.028		−0.037		−0.038		−0.070	us
NG*/NG			3/5		3/5		3/5		4/5
Real to shortest		**Gap** G_k		**Gap** G_k		**Gap** G_k		**Gap** G_k	
ρ_{NOD}	−	−0.251	es*	−0.226	es*	−0.242	es*	−0.205	es*
N_{MAN}	−	0.019		0.027		0.029		0.063	us
ρ_{LEFT}	−	−0.080	es*	−0.078	es*	−0.079	es*	−0.099	es*
ρ_{RIGHT}	−	−0.057		−0.054		−0.056	es*	−0.067	es*
$\rho_{STRAIGHT}$	+	−0.059		−0.051		−0.058		0.006	
ρ_{TURN}	−	−0.137	es*	−0.132	es*	−0.134	es*	−0.166	es*
ρ_{NOTL}	−	−0.269	es*	−0.251	es*	−0.268	es*	−0.271	es*
ρ_{TL}	+	0.019		0.025		0.026		0.066	
ρ_{PRIOCH}	−	−0.093	es*	−0.110	es*	−0.116	es*	−0.185	es*
NG*/NG			5/5		5/5		6/6		6/7
Gains		**Gap** G_k		**Gap** G_k		**Gap** G_k		**Gap** G_k	
ρ_{NOD}	+	0.118	es*	0.088		0.097	es*	0.029	
N_{MAN}	+	0.009		0.006		0.007		−0.004	
ρ_{LEFT}	+	0.019		0.017		0.018		0.015	
ρ_{RIGHT}	+	0.006		0.004		0.005		−0.002	
$\rho_{STRAIGHT}$	−	0.057		0.045		0.052		−0.016	
ρ_{TURN}	+	0.023		0.021		0.021		0.015	
ρ_{NOTL}	+	0.134	es*	0.106	es*	0.117	es*	0.072	
ρ_{TL}	−	−0.015		−0.017		−0.017		−0.039	es*
ρ_{PRIOCH}	+	0.056	es*	0.068	es*	0.073	es*	0.120	es*
NG*/NG			3/3		2/2		3/3		2/2
Losses		**Gap** G_k		**Gap** G_k		**Gap** G_k		**Gap** G_k	
ρ_{NOD}	−	−0.132	es*	−0.138	es*	−0.144	es*	−0.176	es*
N_{MAN}	−	0.028		0.033		0.035	us	0.059	us
ρ_{LEFT}	−	−0.061	es*	−0.060	es*	−0.061	es*	−0.084	es*
ρ_{RIGHT}	−	−0.051	es*	−0.050	es*	−0.050	es*	−0.069	es*
$\rho_{STRAIGHT}$	+	−0.002		−0.006		−0.006		−0.010	
ρ_{TURN}	−	−0.114	es*	−0.111	es*	−0.113	es*	−0.151	es*
ρ_{NOTL}	−	−0.135	es*	−0.144	es*	−0.151	es*	−0.199	es*
ρ_{TL}	+	0.004		0.008		0.009		0.027	
ρ_{PRIOCH}	−	−0.037	es*	−0.042	es*	−0.043	es*	−0.065	es*
NG*/NG			6/6		6/6		6/7		6/7
Total NG*/NG		17/19		16/18		18/21		18/21	

Legend: Different choices of $\boldsymbol{\alpha}$, correspondent parameters' vector $\widehat{\boldsymbol{\beta}}_k^*$, and working hypothesis. The specification (3) does not correspond to any choice of $\boldsymbol{\alpha}$; the coefficients' vector is estimated over the pooled dataset. In this case, the discrimination is expected to be slightly more against female cyclists because the male cyclist group has slightly more observations and, thus, weight.

Specification	Value of $\boldsymbol{\alpha}$	Values of $\widehat{\boldsymbol{\beta}}_k^*$	Estimated gender gap, G_k	Expected direction of discrimination
(1)	0	$\widehat{\boldsymbol{\beta}}_k^1$	$\overline{\boldsymbol{X}}^{0'}(\widehat{\boldsymbol{\beta}}_k^1 - \widehat{\boldsymbol{\beta}}_k^0)$.	Only against female cyclists
(2)	–	$\widehat{\boldsymbol{\beta}}_k^p$	$\overline{\boldsymbol{X}}^{1'}(\widehat{\boldsymbol{\beta}}_k^1 - \widehat{\boldsymbol{\beta}}_k^p) + \overline{\boldsymbol{X}}^{0'}(\widehat{\boldsymbol{\beta}}_k^p - \widehat{\boldsymbol{\beta}}_k^0)$	Slightly more against female cyclists

(*continued on next page*)

(*continued*)

Specification	Value of α	Values of $\widehat{\beta}_k^*$	Estimated gender gap, G_k	Expected direction of discrimination
(3)	0.5	$(\widehat{\beta}_k^0 + \widehat{\beta}_k^1)/2 = \widehat{\beta}_k^{0.5}$	$\overline{X}^{1'}(\widehat{\beta}_k^1 - \widehat{\beta}_k^{0.5}) + \overline{X}^{0'}(\widehat{\beta}_k^{0.5} - \widehat{\beta}_k^0)$.	No discrimination
(4)	1	$\widehat{\beta}_k^0$	$\overline{X}^{1'}(\widehat{\beta}_k^1 - \widehat{\beta}_k^0)$.	Only against male cyclists

Table A5

Decompositions by class of average speed and length: females versus male cyclists

			Length ≤2500 m						Length >2500 m					
			Speed ≤3.5 m/s			Speed >3.5 m/s			Speed ≤3.5 m/s			Speed >3.5 m/s		
K	Route attribute	Exp. sign	Gender gap, *G_k*	Eval	% Effect	Gender gap, *G_k*	Eval	% Effect	Gender gap, *G_k*	Eval	% effect	Gender Gap, *G_k*	Eval	% Effect
1	ρ_{NOD}	-	0.033			-0–044			−0.299			0.284		
2	N_{MAN}	-	−0.030			−0.043			−0.062			0.124	us	
3	ρ_{LEFT}	-	−0.117			−0.150			0.013			0.012		
4	ρ_{RIGHT}	–	−0.155	es*		−0.192	es*		−0.011			0.002		
5	$\rho_{STRAIGHT}$	+	0.323	es*		0.310			−0.235			0.219	es*	
6	ρ_{TURN}	-	−0.272	es*		−0.342	es*		0.003			0.014		
7	ρ_{NOTL}	-	0.078			0.067			−0.136			0.330	us	
8	ρ_{TL}	+	−0.045			−0.111			−0.092			−0.081		
9	ρ_{PRIOCH}	-	−0.051			−0.067			−0.067			−0.009		
	NG/NG*			**3/3**			**2/2**			**0/0**			**1/3**	
k	Real to shortest	Exp. sign	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gendergap, *G_k***	Eval	% Effect	**Gendergap, *G_k***	Eval	% Effect
1	ρ_{NOD}	-	−0.347	es*		−0.432	es*		−0.136			−0.081		
2	N_{MAN}	-	0.035			0.015			−0.075			0.075		
3	ρ_{LEFT}	-	−0.106			−0.046			−0.136	es*		−0.028		
4	ρ_{RIGHT}	–	−0.039			−0.047			−0.107			−0.045		
5	$\rho_{STRAIGHT}$	+	−0.142			−0.210			0.089			−0.051		
6	ρ_{TURN}	-	−0.144			−0.093			−0.243	es*		−0.073		
7	ρ_{NOTL}	-	−0.358	es*		−0.501	es*		−0.153			−0.104		
8	ρ_{TL}	+	0.011			0.069			0.017			0.023		
9	ρ_{PRIOCH}	-	−0.253	es*		−0.083			0.019			−0.075		
	NG/NG*			**3/3**			**2/2**			**2/2**			**0/0**	
k	Gains	**Exp. sign**	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect
1	ρ_{NOD}	+	0.200	es*		0.205			0.035			−0.034		
2	N_{MAN}	+	0.005			0.002			0.049			−0.001		
3	ρ_{LEFT}	+	0.047			−0.015			0.023			−0.003		
4	ρ_{RIGHT}	+	0.020			−0.075			0.053	es*		0.003		
5	$\rho_{STRAIGHT}$	–	0.136			0.108			−0.079			0.059		
6	ρ_{TURN}	+	0.052			−0.079			0.066	es*		0.008		
7	ρ_{NOTL}	+	0.205	es*		0.280	es*		0.015			−0.002		
8	ρ_{TL}	–	0.042			−0.042			−0.028			−0.046		
9	ρ_{PRIOCH}	+	0.181	es*		0.054			−0.058			0.037		
	NG/NG*			**3/3**			**1/1**			**2/2**			**0/0**	
k	Losses	**Exp. sign**	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect	**Gender gap, *G_k***	Eval	% Effect
1	ρ_{NOD}	-	−0.148			−0.227	es*		−0.101			−0.116	es*	
2	N_{MAN}	-	0.040			0.017			−0.025			0.074	us	
3	ρ_{LEFT}	-	−0.059			−0.061			−0.112	es*		−0.031		
4	ρ_{RIGHT}	–	−0.019			−0.122	es*		−0.055			−0.042		
5	$\rho_{STRAIGHT}$	+	−0.005			−0.102			0.010			0.008		
6	ρ_{TURN}	-	−0.092			−0.171			−0.178	es*		−0.065		
7	ρ_{NOTL}	-	−0.154	es*		−0.222	es*		−0.138			−0.106	es*	
8	ρ_{TL}	+	0.054			0.028			−0.011			−0.023		
9	ρ_{PRIOCH}	-	−0.072			−0.029			−0.039			−0.038		
	NG/NG*			**1/1**			**3/3**			**2/2**			**2/3**	
	Total NG/NG*			10/10			8/8			6/6			3/6	

References

Abasahl, F., Kelarestaghi, K.B., Ermagun, A., 2018. Gender gap generators for bicycle mode choice in Baltimore college campuses. Travel Behaviour and Society 11, 78–85.

AitBihiOuali, L., Klingen, J., 2022. Inclusive roads in NYC: gender differences in responses to cycling infrastructure. Cities 127, 103719.

Akar, G., Clifton, K.J., 2009. Influence of individual perceptions and bicycle infrastructure on decision to bike. In: Transportation Research Record: Journal of the Transportation Research Board: No. 2140. Transportation Research Board of the National Academies, Washington, D.C., pp. 165–172

Akar, G., Fischer, N., Namgung, M., 2013. Bicycling choice and gender case study: the Ohio State University. International Journal of Sustainable Transport 7 (5), 347–365.
Aldred, R., Elliott, B., Woodcock, J., Goodman, A., 2017. Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age. Transport Rev. 37–1, 29–55.
Arabikhan, F., Postorino, M.N., Dupont-Kieffer, A., Gegov, A., 2016. Gender-based analysis of zones of tolerance for transit service quality considering intelligent transportation systems. Transport. Res. Rec. 2541, 73–80.
Baker, L., 2009. Shifting gears. Sci. Am. 301 (4), 28–29.
Blais, A.R., Weber, E.U., 2001. Domain-specificity and gender differences in decision making risk. Decision and Policy 6 (1), 47–69.
Blinder, A.S., 1973. Wage discrimination: reduced form and structural estimates. J. Hum. Resour. 8, 436–455.
Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transport. Res. Pol. Pract. 46 (10), 1730–1740.
Buehler, R., Pucher, J., Merom, D., Bauman, A., 2011. Active travel in Germany and the U.S.: contributions of daily walking and cycling to physical activity. Am. J. Prev. Med. 41 (3), 241–250.
Carroll, J., Brazil, W., Morando, B., Denny, E., 2020. What drives the gender-cycling-gap? Census analysis from Ireland. Transport Pol. 97, 95–102.
Cherchi, E., Hensher, D.A., 2015. Workshop synthesis: stated preference surveys and experimental design, an audit of the journey so far and future research perspectives. Transport. Res. Procedia 11, 154–164.
Cotton, J., 1988. On the decomposition of wage differentials. Rev. Econ. Stat. 70 (2), 236–243.
Dill, J., Gliebe, J., 2008. Understanding and Measuring Bicycling Behavior: a Focus on Travel Time and Route Choice. Oregon Transportation Research and Education Consortium, Portland, OR.
Dill, J., Goddard, T., Monsere, C.M., McNeil, N., 2015. Can Protected Bike Lanes Help Close the Gender Gap in Cycling? Lessons from Five Cities. Annual Meeting of the Transportation Research Board, Washington, D.C. Presented at the 94th.
Echiburú, T., Hurtubia, R., Muñoz, J.C., 2021. The role of perceived satisfaction and the built environment on the frequency of cycle-commuting. Journal of Transport and Land Use 14 (1), 171–196.
Emond, C.R., Tang, W., Handy, S.L., 2009. Explaining gender differences in bicycling behavior. Transport. Res. Rec. 2125, 16–25.
Garrard, J., 2003. Healthy revolutions: promoting cycling among women. Health Promot. J. Aust. 14, 213–215.
Garrard, J., Rose, G., Lo, S.K., 2008. Promoting transportation cycling for women: the role of bicycle infrastructure. Prev. Med. 46, 55–59.
Good, I.J., 1992. The Bayes/Non-Bayes compromise: a brief review. J. Am. Stat. Assoc. 87 (419), 597–606.
Grudgings, N., Hughes, S., Hagen-Zanker, A., 2021. The comparison and interaction of age and gender effects on cycling mode-share: an analysis of commuting in England and Wales. J. Transport Health 20, 101004.
Harris, C.R., Jenkins, M., Glaser, D., 2006. Gender differences in risk assessment: why do women take fewer risks than men? Judgement and Decision Making 1, 48–63.
Heesch, K.C., Sahlqvist, S., Garrard, J., 2012. Gender differences in recreational and transport cycling: a cross-sectional mixed-methods comparison of cycling patterns, motivators, and constraints. Int. J. Behav. Nutr. Phys. Activ. 9, 106.
Heinen, E., Maat, K., van Wee, B., 2013. The effect of work-related factors on the bicycle commute mode choice in The Netherlands. Transportation 40, 23–43.
Istat, 2017 (in Italian). https://www.istat.it/it/files//2018/11/Report-mobilità-sostenibile.pdf. (Accessed 7 February 2022).
Istat, 2020. Gli spostamenti sul territorio prima del Covid-19. https://www.istat.it/it/files//2020/05/spostamenti-sul-territorio_2019.pdf (in Italian). (Accessed 7 February 2022).
ITDP, 2022. Cycling's gender gap: breaking the cycle of inequality. https://www.itdp.org/2022/07/06/cyclings-gender-gap/. (Accessed 15 October 2022).
Jensen, A.F., Palao, A.F., Rasmussen, T.K., Nielsen, O.A., 2019. Using crowd source data in bicycle route choice modeling. In: Proceedings from the Annual Transport Conference at Aalborg University, vol. 26, 1.
Krizek, K.J., Johnson, P.J., Tilahun, N., 2005. Gender differences in bicycling behavior and facility preferences. Research on Women's Issues in Transportation: Report of a Conference 2, 31–40.
Lusk, A.C., Wen, X., Zhou, L., 2014. Gender and used/preferred differences of bicycle routes, parking, intersection signals, and bicycle type: professional middle class preferences in Hangzhou, China. J. Transport Health 1, 124–133.
Matsuda, F., Ikegami, T., Kishida, K., Kimotsuki, K., 2000. Effects of gender, age and experience on bicycle riding behavior. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications, p. 417.
Metropolitana di Bologna, Città, 2017. Bella Mossa Campaign (in Italian). https://www.cittametropolitana.bo.it/portale/Home/Archivio_news/Bella_Mossa_chi_si_muove_bene_si_premia_. (Accessed 7 February 2022). Accessed.
Moudon, A., Lee, C., Cheadle, A., Collier, C., Johnson, D., Schmid, T., Weather, R., 2005. Cycling and the built environment, a US Perspective. Transport. Res. Transport Environ. 10 (3), 245–261.
Municipality of Bologna, 2020. Il monitoraggio dei flussi di ciclisti su alcune sezioni della rete ciclabile di Bologna (in Italian). http://www.comune.bologna.it/media/files/report_flussi_ciclabili_2019.pdf. (Accessed 7 February 2022).
Nelson, A.C., Allen, D., 1997. If you build them, commuters will use them: association between bicycle facilities and bicycle commuting. Transport. Res. Rec. 1578, 79–83.
Nelson, T., Roy, A., Ferster, C., Fischer, J., Brum-Bastos, V., Laberee, K., Yu, H., Winters, M., 2021. Generalized model for mapping bicycle ridership with crowdsourced data. Transport. Res. C Emerg. Technol. 125.
Neumark, D., 1988. Employers' discriminatory behavior and the estimation of wage discrimination. J. Hum. Resour. 23 (3), 279–295.
Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. Int. Econ. Rev. 14 (3), 693–709.
Picasso, E., Postorino, M.N., Bonoli-Escobar, M., Stewart-Harris, M., 2020. Car-sharing vs bike-sharing: a choice experiment to understand young people behaviour. Transport Pol. 97, 121–128.
Poliziani, C., Rupi, F., Mbuga, F., Schweizer, J., Tortora, C., 2020. Categorizing three active cyclist typologies by exploring patterns on a multitude of GPS crowdsourced data attributes. Research in Transportation Business Management 40, 100572.
Poliziani, C., Rupi, F., Schweizer, J., Postorino, M.N., Nocera, S., 2022. Modeling cyclist behavior using entropy and GPS data. International Journal of Sustainable Transportation 1556–8318. https://doi.org/10.1080/15568318.2022.2079446.
Prati, G., Fraboni, F., De Angelis, M., Pietrantoni, L., Johnson, D., Shires, J., 2019. Gender differences in cycling patterns and attitudes towards cycling in a sample of European regular cyclists. J. Transport Geogr. 78, 1–7.
Pucher, J., Buehler, R., 2008. Making cycling irresistible: lessons from The Netherlands, Denmark, and Germany. Transport Rev. 28 (4), 495–528.
Pucher, J., Buehler, R., Merom, D., Bauman, A., 2011a. Walking and cycling in the United States, 2001–2009: evidence from the national household travel surveys. Am. J. Publ. Health 11 (S1), S310–S317.
Pucher, J., Garrard, J., Greaves, S., 2011b. Cycling down under: a comparative analysis of bicycling trends and policies in Sydney and Melbourne. J. Transport Geogr. 19 (2), 332–345.
Reimers, C.W., 1983. Labor market discrimination against hispanic and black men. Rev. Econ. Stat. 65 (4), 570–579.
Rupi, F., Poliziani, C., Schweizer, J., 2019. Data-driven bicycle network analysis based on traditional counting methods and GPS traces from smartphone. ISPRS Int. J. Geo-Inf. 8 (8), 322.
Rupi, F., Poliziani, C., Schweizer, J., 2020. Analysing the dynamic performances of a bicycle network with a temporal analysis of GPS traces. Case Studies on Transport Policy 8 (3), 770–777.
Schoner, J., Lindsey, G., Levinson, D., 2015. Factors associated with the gender gap in bicycling over time. In: Presented at 94th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.
Schweizer, J., Rupi, F., 2014. Performance evaluation of extreme bicycle scenarios. Procedia Social and Behavioral Sciences 111, 508–517.
Schweizer, J., Bernardi, S., Rupi, F., 2016. Map-matching algorithm applied to bicycle global positioning system traces in Bologna. IET Intell. Transp. Syst. 10 (4), 244–250.
Sener, I.N., Eluru, N., Bhat, C.R., 2009. An analysis of bicycle route choice preferences in Texas, US. Transportation 36, 511–539.
Shaw, C., Russell, M., Keall, M., MacBride-Stewart, S., Wild, K., Reeves, D., Bentley, R., Woodward, A., 2020. Beyond the bicycle: seeing the context of the gender gap in cycling. J. Transport Health 18, 100871.
Special Eurobarometer 422, 2014. Quality of transport. TNS opinion & social - directorate-general for mobility and transport (DG MOVE). https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_422a_en.pdf. (Accessed 7 February 2022). Accessed.
Stanley, T.D., Jarrell, S.B., 1998. Gender wage discrimination bias? A metaregression analysis. J. Hum. Resour. 33, 947–973.
Stinson, M., Bhat, C., 2004. Frequency of bicycle commuting: internet-based survey analysis. Transport. Res. Rec. 1878 (1), 122–130.
Sustrans, 2018. Inclusive City Cycling Women: Reducing the Gender Gap. https://www.sustrans.org.uk/media/2930/2930.pdf. (Accessed 7 February 2022). Accessed.
Tilahun, N.Y., Levinson, D.M., Krizek, K.J., 2007. Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Transport. Res. Pol. Pract. 41, 287–301.
Tukey, J., 1991. The philosophy of multiple comparisons. Statistics Science 6 (1), 100–116.
Twaddle, H., Hall, F., Bracic, B., 2010. Latent bicycle commuting demand and effects of gender on commuter cycling and accident rates. Transport. Res. Rec. 2190, 28–36.
U.S. Census Bureau, 2019. https://www.census.gov/library/stories/2019/05/younger-workers-in-cities-more-likely-to-bike-to-work.html. (Accessed 7 February 2022). Accessed.
Weber, E.U., Blais, A., Betz, E.N., 2002. A domain specific risk-attitude scale: measuring risk perceptions and risk behaviors. J. Behav. Decis. Making 15, 263–290.
Weichselbaumer, D., Winter-Ebmer, R., 2005. A meta-analysis of the international gender wage gap. J. Econ. Surv. 19, 479–511.
Winters, M., Zanotto, M., 2017. Gender trends in cycling over time: an observational study in Vancouver, British Columbia. J. Transport Health 5, S37–S38.
Winters, M., Friesen, M.C., Koehoorn, M., Teschke, K., 2007. Utilitarian bicycling: a multilevel analysis of climate and personal influences. Am. J. Prev. Med. 32 (1), 52–58.
Winters, M., Davidson, G., Kao, D., Teschke, K., 2011. Motivators and deterrents of bicycling: comparing influences on decisions to ride. Transportation 38 (1), 153–168.