urban science

MDPI

Article

Determinants of Fare Evasion in Urban Bus Lines: Case Study of a Large Database Considering Spatial Components

Susana Freiria [1,2,*] and Nuno Sousa [3,4]

1 Department of Environment and Planning, Universidade de Aveiro, 3810-193 Aveiro, Portugal
2 Department of Civil Engineering, University of Coimbra, CITTA, Polo II, 3030-788 Coimbra, Portugal
3 Departamento de Ciências e Tecnologia, Universidade Aberta, 1269-001 Lisbon, Portugal; nsousa@uab.pt
4 INESCC-Coimbra—Institute for Systems Engineering and Computers at Coimbra, 3030-790 Coimbra, Portugal
* Correspondence: susana.freiria@ua.pt

Abstract: This article presents a large case study of fare evasion on bus lines in the city of Lisbon, Portugal, a common problem in dense urban areas. Focus is put on geographic factors, and an analysis is carried out using a generalized spatial two-step least-squares regression (GS2SLS). The large database, spanning one year of fare evasion reports, made it possible to segregate the analysis according to type of day (workday or weekend) and time period (rush hours, nighttime, etc.). Results show that indeed the type of day and time period lead to considerable differences between the seven models analyzed. It was found that the number of inspection actions is the strongest predictor of evasion, with geographic factors also playing a role in predicting fare evasion. Consideration of this spatial component made it possible to find moderate evidence for dissuasive effects of inspection actions in some models and of pockets of evasive tendencies in other models, which appear in the statistical error term. Interestingly, and contrary to other studies, age was found to have almost no influence on the propensity to evade fares. From a managerial point of view, this study highlights the importance of running inspection actions systematically and closely monitoring their outcomes. From a more theoretical standpoint, it suggests further research on geographic factors is needed to fully understand spatial patterns of evasion.

Keywords: fare evasion; public transport; GS2SLS regression; urban transport governance

check for updates

Academic Editor: Chia-Yuan Yu

Received: 1 May 2025
Revised: 7 June 2025
Accepted: 11 June 2025
Published: 18 June 2025

Citation: Freiria, S.; Sousa, N. Determinants of Fare Evasion in Urban Bus Lines: Case Study of a Large Database Considering Spatial Components. *Urban Sci.* **2025**, *9*, 231. https://doi.org/10.3390/urbansci9060231

1. Introduction

Fare evasion represents a global problem with consequences at several levels. At an economic level, it has repercussions on the revenue generated, price of the ticket, level of government support, and on the image of public transport companies. At a social level, it generates feelings of insecurity and social injustice. It has also been demonstrated that if fare evasion reaches a critical level, passengers who never evaded fares start doing so, frustrated by the high number of evasions they observe [1–3], leading to a snowball effect. Therefore, public transport companies must spend financial and human resources to mitigate fare evasion.

One of the fundamental questions that come up is to identify who the evaders might be and why they evade, i.e., what factors might influence the propensity to evade fares. Surveys have been carried out to identify the socio-demographic characteristics as well as the motivations of fare evaders [4–6]. However, there is no consensus among the authors on a standard evader profile: one study [7] concluded that young male passengers have a higher probability of committing fare evasion, but other studies [1,8] found that other variables are more significant, such the time of the day [1,7]. Most samples are usually quite small, thus not very representative, and many inquire passengers during rush hour, which may not be the best strategy to profile the potential evaders. More recently, some authors suggested that a solution might be to use massive data from smartcard transactions [8]. Nevertheless, this data source still has some limitations because not all the individuals who fail to validate their tickets are evaders [9,10].

The main goal of this work is to address the above-mentioned literature gap by identifying more accurately the factors involved in fare evasion, including geographic ones. The methodology proposed here is different from previous approaches in that, instead of using surveys, the approach proposed here relies on inspective operational records at the bus stop scale, combined with census data, which does not require directly surveying passengers. Census data provide socioeconomic characteristics of the individuals living in the surrounding areas of bus stops, and this research makes use of those data to uncover to what extent those characteristics can explain the detected fare evasion, adding geographical considerations to the role of possible explanations to fare evasion. Such a fine-grained analysis can potentially contribute to, e.g., allocating field inspectors more efficiently or influencing bus network design algorithms (e.g., [3,11]). Previous studies usually present results at a global, city-wide scale [12–14], neglecting the spatial component, i.e., possible differences in fare evasion patterns that depend on geographic factors. To fill this gap, the methodological approach is based on generalized spatial two-stage least squares (GS2SLS), a statistical model which can capture spatial dependence.

The case study took place in Lisbon, the capital city of Portugal, and data refer to evasion in Carris, the main collective public transport bus company for passengers, in 2019. During that year, Carris transported approximately 44 M passengers and carried out 1.84 M inspection actions. The size of the sample is thus considerably larger than those used in previous works. This large database made it possible to identify trends in the results that, upon analysis, shed light into previous results in the literature, added new insights of considerable relevance from a managerial point of view, and, as will be seen, suggests future research avenues to deepen spatial considerations.

2. Literature Review

In their extensive review, the authors of [13] argued that the topic of fare evasion covers very diverse areas, ranging from ticketing infrastructures to the sociodemographic characteristics of evaders. Those authors identified five main areas of research: #1 fare evader-oriented studies, #2 criminology, #3 economics, #4 technological innovations, and #5 operational research. The main goal of evader-oriented studies is to identify the key characteristics and motivations of fare evaders [15]. With respect to criminology, fare evasion can be considered a crime and may trigger other crimes [15]. Understanding this link is a major objective of this subfield of criminology [13,15]. Economically, the interest is mainly to study the impact of fare evasion in the balance sheet of transport companies. Technological innovations have also deserved attention mainly due to applications to ticketing systems. Finally, fare evasion gives rise to operational research problems, such as, e.g., the development of optimization models to schedule efficiently inspector teams [16]. The present research is positioned in the first field of research mentioned in [13], given that it aims at defining an evader profile and investigating what factors and geographic patterns may exist for evasion. The remaining of this section concentrates on this aspect.

Over the years, work has been carried out on fare evasion in public transport in different countries, e.g., Australia [17,18], Belgium [19], Chile [12], USA [20], France [5,21], and Italy [4,22]. Despite the different geographic quadrants and difficulties found, it was possible to identify some common characteristics among passengers who commit fare evasion.

In a large study [4], 16,000 passengers of 12 Italian transport companies were interviewed, authors having concluded that younger, unemployed, and foreign passengers were more likely to commit fare evasion [4,23,24]. In [22], different methodologies from those of [4] were used, but similar conclusions were found. In addition to these studies developed in Italy, a study in Flanders (Belgium) [20] interviewed 636 passengers and also concluded that young males were more likely to commit fare evasion [19]. However, youth does not seem to point in the same direction everywhere; for example, in [18], fare evasion in the Melbourne Metropolitan Area (Australia) was analysed, and it was found that males under the age of 25 were actually less likely to commit fare evasion. Recently [7], a new approach was proposed based on a hybrid discrete choice model. By analysing the characteristics of 324 evaders caught in the act of committing a crime in Bogotá (Colombia) [24], the authors assessed the reasons they cited for committing fare evasion. The latent variables were the level of satisfaction with the transport system and the level of self-awareness of the evaders, authors having concluded that the level of self-awareness of evaders is significantly influenced by age; that is, evaders under the age of 29 had a lower level of self-awareness, showing age and evasion might be relating indirectly.

As for the reasons for fare evasion, passengers being unable to pay the ticket certainly plays a role [5,25]. Indeed, in [2,5,20,26], it was found that passengers of below-average income are more likely to commit fare evasion [2]. In this context, the problem no longer arises in terms of the probability of being inspected but of the probability of being fined. Authors of [1,21,27] agree that the greater the certainty of being fined, the lower the propensity to evade. However, other researchers [7] found no relationship between possible punishment and tariff evasion [24]. As shown by a study in Santiago, Chile [28], other factors, such as dissatisfaction with the service and high fares, also played a role in predicting infringements [28]. Indeed, the shortcomings of the bus system led users to develop a negative perception of it, which in turn made fare evasion more socially acceptable.

Another relevant issue is knowing to what extent inspection actions can constitute a deterrent to fare evasion. While studies usually reveal a positive correlation between the number of inspections and the number of fare evasions [29,30], some studies failed to find a direct relationship between the percentage of passengers inspected and the percentage of fare evasion detected [31]. Other studies ([1] and, more recently, [19,32]) found that the probability of being inspected is only seen as a deterrent to fare evasion up to a certain limit.

Analysis of the results obtained by the works mentioned above indicates that it is not possible to completely eradicate tariff evasion, as inspections and fines can only at best mitigate the problem. In fact, using the case of San Francisco (USA) as a reference, Lee [7] concluded that it was not possible to define the profile of the typical intentional evader and that evasion was mainly driven by contextual factors in that case study.

Table 1 summarizes some of the main studies developed in area #1 of [13]. Most of the works are based on interviews to passengers, and the largest sample was composed of 56,746 passengers [9], a sample that is significantly smaller than the sample used in this study. Moreover, most of the studies are focused on a small area of a city or on a bus line, neglecting issues such as spatial autocorrelation. The approach proposed in this research sips on a very large database, covering 1.84 million inspection actions at the bus stop scale, and the methodological approach is based on GS2SLS, a method that considers the spatial component and goes beyond the more usual logistic regression methods that do not consider such dependences (see [33] for an example). Thus, the main contribution of the present research is to provide results on evasion patterns that include a spatial component, resorting to a very large database and study area. Consideration of the spatial component fills a literature gap on the subject, which has so far concentrated on other factors.

Table 1. Studies on socio-demographics and travel determinants of fare evaders.

Authors	Location, City	Type of Survey	Sample Size (Passengers)	Analytical Tool
Abrate et al. [4]	Italy, 12 cities	Intercept (stop) interviews	16,000	Logistic regression
Barabino et al. [23]	Italy, Cagliari	Intercept (on-board) interviews	2177	Logistic regression
Barabino and Salis [13]	Italy, Cagliari	Intercept (on-board) interviews	4404	Logistic regression
Bucciol et al. [22]	Italy, Reggio Emilia	Intercept (stop) interviews	544	Probit regression/correlation analysis
Busco et al. [14]	Chile, Santiago	Intercept (on-board) interviews	503	Factor analysis
Cantillo et al. [10]	Chile, Santiago	Intercept (on-board) interviews	10,559	Binomial Logit Model
Cools et al. [19]	Belgium, Flanders	Web-based questionnaire	638	Logistic regression
Currie and Delbosc [6]	Australia, Melbourne	Web-based questionnaire	1561	Structural equation modelling
Eddy [18]	Australia, Melbourne	Intercept (stop) observations	288	Descriptive statistics
Egu and Bonnel [9]	France, Lyon	Intercept (on-board) interviews	56,746	Descriptive statistics
Delbosc and Currie [15]	Australia, Melbourne	Web-based questionnaire	1561	Two-step cluster analysis
Dai et al. [21]	France, Lyon	Intercept (stop) interviews and lab experiment	279	Descriptive statistics and logistic regression
Gonzalez et al. [28]	Chile, Santiago	Intercept (on-board) interviews	457	*K*-means clustering
Guzman et al. [24]	Colombia, Bogotá	Intercept (stop) interviews	324	Hybrid discrete choice model
Milioti et al. [26]	Greece, Athens	Intercept (stop) interviews	304	Ordinal logistic model

3. Materials and Methods

3.1. Data

Carris [34] provides public road passenger transport services in the Lisbon Metropolitan Area. In 2019, there were 87 routes on offer, with an average length of 19 km (round trip) and 2221 stops, including 1992 (90%) located in the municipality of Lisbon (see Figure 1). In this work, particular emphasis will be placed on stops where at least one inspection action occurred, which corresponds to a total of 1654 stops. Inspection actions involve stopping the bus for checks (in a stop or on-route to the next stop), so each action is tagged to a particular bus stop. The evasion rates oscillated between 2.2% and 12.5% of detected evasions per inspection action, depending on the period of the day and whether it was a working day or weekend. The evasion rate also varied during the year, but this dependence was not studied beyond descriptive statistics because year-to-year comparisons would require data spanning multiple years, which was not available.

The large amount of data made it possible to divide the analysis into weekdays and weekends, and further divide each case into distinct time periods, making it easier to detect time-dependent trends. For the weekdays, these were the morning and evening rush hours (6–10 h, 17–20 h), early and late night (20–0 h, 0–6 h), and daytime off-peak (10–17 h). For the weekend, the division was between late night (0–6 h) and daytime (6–20 h). These time slots were defined based on the operational logic of the inspection teams. The different time periods at the weekend are justified by the smaller universe of available data. The early weekend night period suffered from insufficient data, and, as in [12], it was not possible to analyze it.

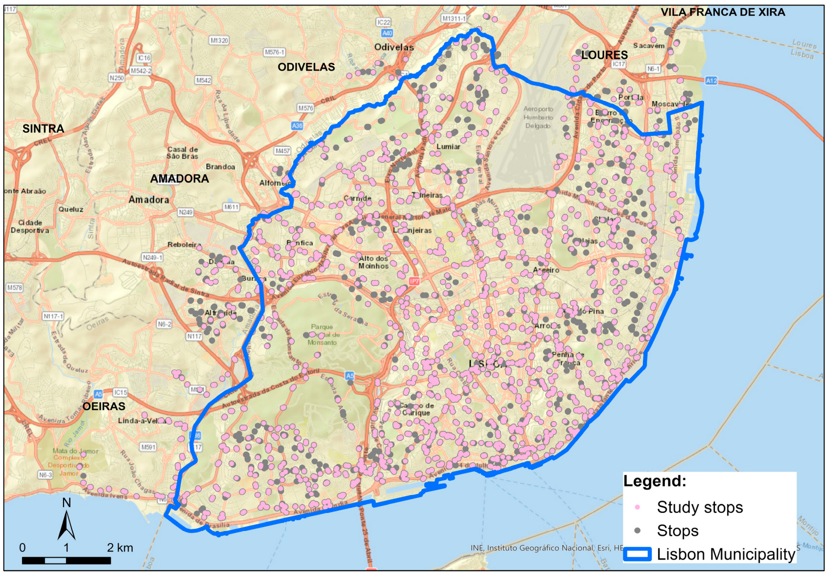


Figure 1. Stops of the Carris bus network.

Figure 2 represents the number of inspection actions per stop. Data analysis reveals that in 53% of stops less than 250 inspection actions were carried out, less than one per day. In 13% of the stops, the number of inspection actions exceeded 2000, with a significant concentration in the busy zone of Avenidas Novas, an electoral ward division in the centre of the city. There was less emphasis on stops outside the municipality of Lisbon, with an average of 337 actions per stop outside the municipality vs. 917 inside its perimeter. This difference merely reflects the fact some bus lines end in a different municipality and are otherwise not meaningful.

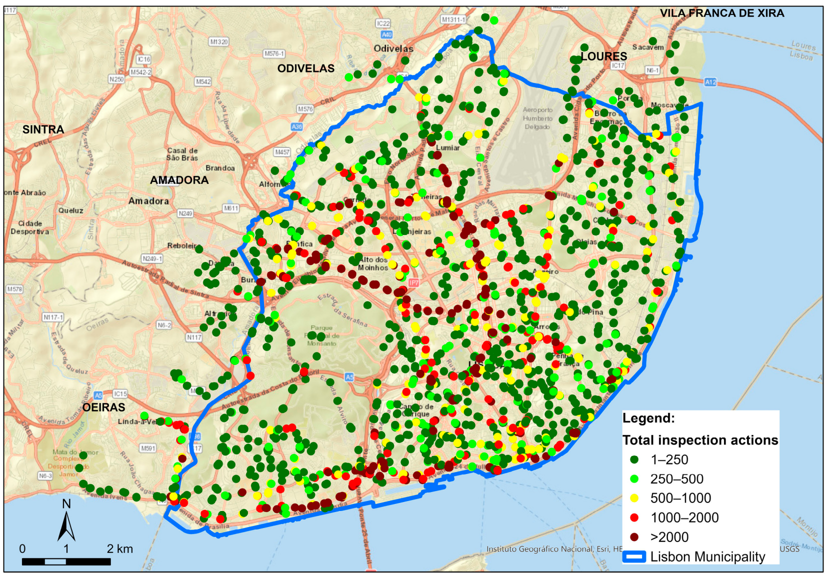
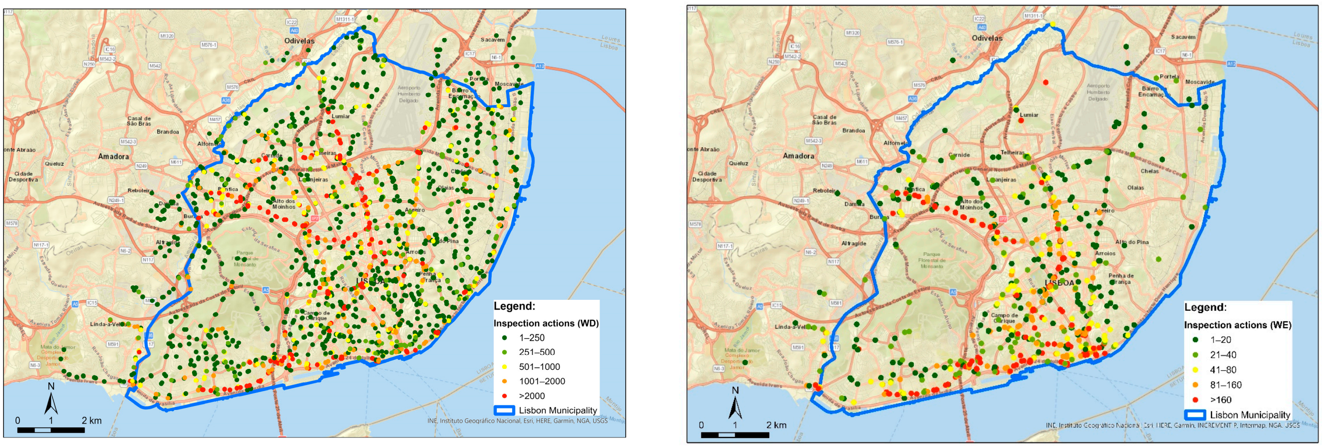


Figure 2. Inspection actions during 2019 in the Carris Network.

The difference between the number of inspection actions carried out on working days and weekends is displayed in Figure 3. Around 5000 inspection actions were carried out per day on working days and 300 inspection actions per day on weekends. This difference in numbers is related to the greater number of passengers and service frequency on working days. The average rate of fare evasion detected at the weekend (3.7 infringements per approach, see Table 2) is, however, higher than the average rate detected during the week (2.7 infringements per approach). In both periods, there was a concentration of inspection actions along the routes with the most bus traffic.



(**a**) Working days (**b**) Weekends

Figure 3. Inspection action during 2019 in the Carris network during (**a**) working days and (**b**) weekends.

Table 2. Case study descriptive statistics on fare evasion.

		Weekdays						Weekend	
	Statistic	**0–6 h**	**6–10 h**	**10–17 h**	**17–20 h**	**20–0 h**	**Total (WD)**	**0–6 h**	**6–20 h**
Inspection actions (IA)	Total	16,347	342,647	646,125	295,531	48,206	1,348,856	9029	62,693
	IA/h	10.9	341.3	367.7	392.5	48.0	223.9	14.5	43.1
Detected evasion (DE)	Total	355	8752	19,159	8027	1716	38,009	1130	2101
	DE/h	0.24	8.72	10.90	10.66	1.71	6.31	1.81	1.44
Detection rate	DE/IA × 100%	2.17	2.55	2.97	2.72	3.56	2.82	12.5	3.4

Table 2 provides global statistics on fare evasion.

The table shows differences between weekdays and weekends, as well as differences throughout the day. Trends of greater propensity for fare evasion at certain times of the day have also been identified by other authors (e.g., [1,2,12]). However, scarcity of data is often highlighted as the reason why this subject has not been explored further, a problem the present study does not have.

On working days, the highest detected fare evasion rate of 3.56% per inspection occurs early night (20–0 h) and coincides with the period of lowest number of inspection actions outside late night—48 per hour. The late-night period (0–6 h) evasion detection rate varied substantially between weekdays and weekends, oscillating between the lowest and highest values (2.17% vs. 12.5%). Greater propensity for fare evasion at night was also reported in [1,35], although those authors did not differentiate between working days and weekends [11]. That difference was considered in [10], and those authors did observe higher evasion rates on weekends than on working days, as well as an increase in evasion trends as the day progressed. That trend also shows in the present research, with a dip in the evening rush hour (17–20 h), which can perhaps be explained by the fact that many workers who travel during this peak hour have monthly passes.

Figure 4 shows the monthly evolution of passengers, inspection actions, and detected fare evasions (i.e., infringements) in 2019 in a normalized 0–1 scale (0 = minimum value, 1 = maximum value). The line patterns suggest some degree of correlation between the three variables, and an analysis reveals the highest $R = 0.63$ between inspection actions and infringements ("you find what you look for"), and the lowest $R = 0.55$ between inspection actions and passengers. This is a moderate degree of correlation that suggests inspection actions are not the only reason infringements are detected. This leaves open the possibility that other factors influence the propensity to evade fares.

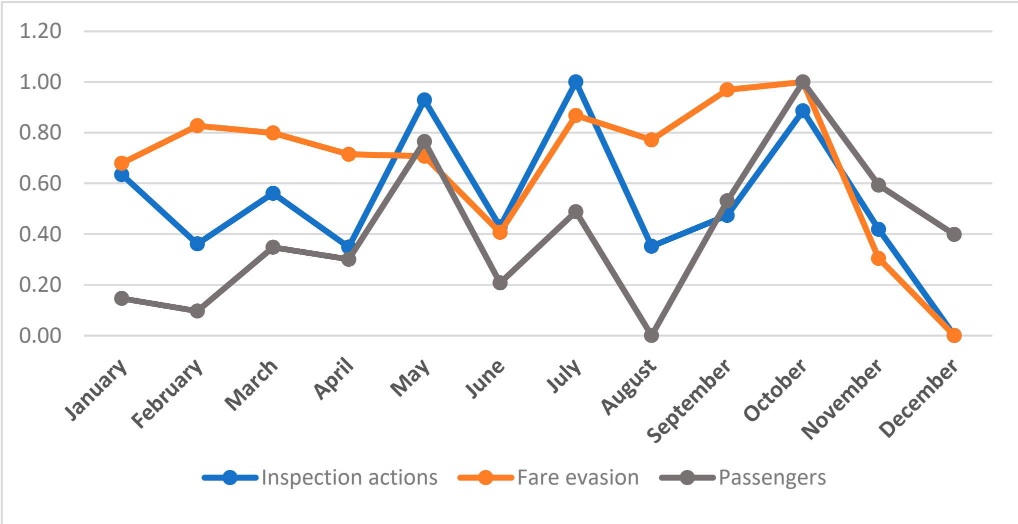


Figure 4. Number of inspection actions, fare evasions, and passengers per month during 2019.

3.2. Model Formulation

Having ascertained that factors other than inspection actions contribute to explain fare evasion tendencies, the next step is to assess what those factors may be. For this purpose, a model is proposed to determine to what extent operational variables and socio-economic characteristics of the individuals living in the surrounding areas of the bus stops have a significant impact on the number of fare evasions.

Working at the scale of the bus stop brings challenges, such as spatial lag and spatial error dependencies. The generalized spatial two-stage least-squares regression model (GS2SLS) considers these two issues, as the method contains a spatially lagged dependent variable and a spatial autocorrelation error term [36,37]. Following those references, the GS2SLS model proposed to examine the impact of socio-economic variables, as well as operational variables on detected fare evasion is defined as follows:

$$D_i = \beta_0 + \sum_k \beta_k X_{ki} + \rho \sum_j W_{ij} D_i + \mu_i, \quad \mu_i = \lambda \sum_j W_{ij} \mu_j + \varepsilon_i \tag{1}$$

where

D_i is the dependent variable; the number of detected infringements in bus stop is i;

X_{ki} is the value of the k-th explanatory variable at bus stop i (see below);

β_k is regression coefficient for the k-th explanatory variable, with β_0 as the model intercept;

ρ is the parameter of spatial autocorrelation;

W_{ij} is the queen contiguity matrix element between bus stops i and j (0 or 1);

μ_i is the error term autocorrelation at bus stop i, with λ being the error autocorrelation parameter and ε_i being the independent and identically distributed statistical error (Note: some authors interchange the role of ρ and λ in their notation.)

Five explanatory variables were considered, described in Table 3.

Table 3. Model variables and description.

Variable	Role	Description	Year	Source
Detection evasion	Dependent	Number of people travelling without a ticket detected by inspection agents at a given stop	2019	Carris [34]
Bus frequency	Explanatory	Number of buses serving each stop at a given period of the day (not used at weekends)	2019	Carris [34]
Age 15–25 (Male)	Explanatory	Number of males between 15 and 25 years old residing in the statistical subsection 1 of the bus stop	2021	INE [38]
POI density	Explanatory	Density of touristic points of interest 2 in the statistical subsection of the bus stop (POI/km^2)	2022	OSM [39]
Inspection actions	Explanatory	Number of inspection actions assigned to a given bus stop	2019	Carris [34]

1. Subdivision of a civil electoral ward with dwellings for 550 to 650 inhabitants. 2. POI as marked by users in OpenStreetMaps.

Bus frequency serves as proxy for the number of passengers [40]. The impact of the number of passengers on fare evasion is not consensual among authors. In some studies [2], a greater number of infringements was identified during rush hours [2], while in others [23], the highest probability for infringement was observed between 12 h and 14 h [22], when the number of passengers is lower than during rush hours.

Young people, here defined as males belonging to the age segment of 15–25 years old, is a factor that several authors [22,23] have found to be associated with a greater propensity for fare evasion [22,23], a trend disputed by other studies. Precise data, as of 2019, were not available due to census being held every 10 years. Note that these data were obtained indirectly, not by inquiring infractors about their age during inspection.

The density of points of interest represents destination attractiveness and was included as a supply-side explanatory variable. Any differences in POI density in pre- and post-COVID-19 years are expected to be small, due to the tourism industry being put on hold in the period in-between (2020–2021).

Finally, inspection actions are a very important variable, as one can only detect what one looks for. The regression coefficients for this variable are interpreted as detected evasion per capita of inspection actions.

Two more variables were considered as potentially significant factors, namely the number of schools within a radius of 500 m from a bus stop and average trip duration of a passenger in the inspected bus. The inclusion of schools was suggested by Carris, as the company suspected of higher evasion near these facilities. However, its use as an explanatory variable was discontinued after results showed it was not significant. Concerning trip duration, greater propensity for infringement on trips lasting less than 15 min was suggested in [21,23]. However, trip duration data were only available from ticket log data, which is related to passengers who travelled legally. Obtaining trip duration data from evaders would require inquiring them directly (and relying on an honest answer), a procedure which was not part of inspection actions. Therefore, this variable was also dropped due to lack of bulk reliable data.

It is important to make a note at this point. Findings described in the Literature Review Section suggest that people with lower income are more likely to evade due to fare price.

A natural variable to include in this study would therefore be average income in the bus stop zone, information that is queried in the census. However, European data protection legislation prevents information from being used in public documents. Authors inquired the statistics bureau of Portugal for its (anonymized) use in academic context, but no permit was given, and so that variable could not be considered.

Another point worth mentioning concerns the trade-offs the dataset entails. Carris inspection actions have as an objective to issue a fine on infractors and do not collect any demographic or social data on them. Inspection actions that collect more data, e.g., sociodemographic data or motivation, would take longer to carry out, lead to smaller datasets, and be more liable to incorrect information. So, in a sense, the dataset used for this research trades-off complementary information on the infraction for a larger sample, subsequently relying on geographic and contextual information (e.g., time of day) for deriving conclusions. Most studies on evasion do the opposite thing. The present study thus complements findings of those studies with evidence gathered from a different systemic approach based on large datasets. Inclusion of spatial endogeneity (via GS2SLS modelling) and two factors with a spatial expression (age, POI density) complement the Carris dataset, and the resulting models are arguably a consistent and coherent way to study fare evasion in context of the available information.

4. Results

Running the model yielded the results in Tables 4 and 5.

Table 4. Fare evasion model estimates—working days.

Detected Infringements	0–6 h	6–10 h	10–17 h	17–20 h	20–0 h
Intercept	1.645	0.276	0.546	1.724 **	0.900
Age 15–25	0.046	0.009	0.012	−0.013	0.034
Rho W	0.012	0.038	−0.20	−0.110 ***	0.023
Inspection actions	0.0124 ****	0.0234 ****	0.0236 ****	0.0268 ****	0.0390 ****
Bus frequency	0.066	0.024	0.336 ***	0.007	−0.286 ***
POI density	−0.351	−0.0979 ****	−0.017	−0.144 ***	−0.073
Lambda W	−0.244	0.039	0.155 ****	0.073	0.208 **
Pseudo $R2$/spatial $R2$	31%/32%	82%/82%	77%/78%	64%/64%	72%/71%

1 Signif. codes: **** 0.001–0.01; *** 0.001–0.01; ** 0.01–0.05. No marking code: 0.1–1.

Table 5. Fare evasion model estimates—weekend.

Detected Infringements	0–6 h	6–20 h
Intercept	0.093	−1.095
Age 15–25	−0.265 **	0.016
Rho W	0.121 *	−0.076
Inspection actions	0.133 ****	0.0330 ****
Bus frequency	0.205	0.004
POI density	0.035	0.069
Lambda W	−0.175	0.031
Pseudo $R2$/spatial $R2$	88%/88%	34%/34%

1 Signif. codes: **** 0.001–0.01; ** 0.01–0.05; * 0.05–0.1. No marking code: 0.1–1.

Looking at the tables, the following trends can be identified:

1. The five derived models exhibit similarities across the day, but the statistical significance of some explanatory variables varies considerably depending on the period of day, confirming that fare evasion is indeed influenced by the time factor. The extent to which the period of the day influences evasion is best seen from Table 2, whose

results confirm higher evasion rates during 10–17 h off-peak, as noted in [22], both in absolute terms and detection rates, placing this period as likely more prone to evasion than rush hours. What Table 4 also shows is that the structure of fare evasion tendencies also varies with the period of the day, as explained below.

2. The intercept term is mostly not relevant. This term represents a baseline tendency for infringement. The fact that such tendencies did not show up can perhaps be explained by their absorption into the explanatory power of the inspection actions. It was only for weekdays 10–17 h that a significant intercept was found, of 1.724 evaders per inspected bus. It is tempting to interpret this as evidence of a structural nature of fare evasion in the 10–17 h off-peak period, i.e., that some evasion is likely to exist, even if no inspections were carried out. It is worth noting that the alternative of running the model with infringements per inspective action as dependent variable does yield significant intercepts, which is expectable as they would then have the interpretation of baseline infringement per inspection action.
3. The age appears not to be relevant, contrary to findings by other authors. In the only case where the model did find a tendency, weekends 0–6 h, that tendency was that young males were less likely to infringe, in line with the findings of [18] and opposite to those of [19,22,23]. As the large database used in this research decreases the likelihood of a statistical type II error (false negative), a case can be made for young age not being a predictor of evasion in general. At most young age can produce local, city-scale effects and are unlikely to be significant as a global, worldwide trend.
4. The dependent variable spatial lag has an influence in some time periods, suggesting that some endogeneity in fare evasion at those periods. The negative coefficient on weekdays 17–20 h is perhaps somewhat unexpected, as it suggests anticorrelation: stops with high infringement detection tend to pair with those with low detection. This hints at inspection actions having a stronger deterrent effect at this period, which is a rush hour (many passengers) and has the highest number of inspections per hour (Table 2), leading to a higher visibility of inspections. Passengers who take the inspected bus lines are likely to have witnessed those inspections and refrain from evading, leading to less detection at the nearby stops. On the other hand, the mild positive correlation at weekends late night (0–6 h) suggests a contagion in infringements, i.e., "if they skip the ticket, so will I".
5. Inspective actions have by far the highest explicative power. This was expected, as one can only detect what one looks for, and is in line with findings of [29,30], who advocate for intensifying inspection actions as the best way to fight fare evasion offences. At this point, it is important to note that, as Figure 2 shows, there is unevenness in the distribution of inspection actions, with some stops being more intensively scrutinized than others. It is because of this imbalance that detected evasion must be controlled for the number of inspection actions. Had the inspection actions been evenly distributed, the number of actions would have very little predictive power, if any at all. But, as will be seen, the high explicative power of such actions does not preclude other statistically significant trends for detected evasion. Returning to the results, detection rates are consistent during the day, with regression coefficients revealing about 0.02–0.03 detections per action, and rising during the night, especially on weekends, climbing to 0.04–0.13 detections, possibly revealing infringers trust that they will not be inspected due to the off-hours. An exception is late night weekdays (0–6 h), which have the lowest infringement rates with about 0.01 detections per action. This might be due to people using buses at these hours on weekdays being mostly early workers, who usually have public transport passes. Note that for weekdays 10–17 h, the baseline tendency adds to the effect of inspective actions, yielding the highest

absolute number of detected infringements. This was the only case where variance could be significantly split into an explanation of baseline plus inspection actions.

6. Bus frequency has a small effect. Positive at weekdays 10–17 h, a possible explanation for it might be the temptation to skip the ticket believing that the inspective workforce, which is limited, cannot be on all buses at the same time and so it may prefer the rush hours, losing grip in the period in-between. At early night, the negative effect may come from a fear to be sent off the bus from infringement at an inconvenient time when other busses are scarce and the surroundings potentially unsafe. The daytime frequency effect may be worth considering in the context of frequency-setting algorithms, such as, e.g., [41].
7. POI density has a negative effect on infringement detection at rush hours, suggesting that some people that head to POI-dense locations, perhaps the tourists or workers at the POI, have a lower tendency to infringe. This point seems worthy of further research.
8. The error spatial lag has positive-signed significance during the 10–17 h and 20–0 h weekday periods, suggesting increased evasion trends in some spatially clustered bus stops. Since the error term captures unknown factors, this clustering may be due to geographic variables that influence evasion but do not feature in the models.
9. Finally, concerning model quality ($R2$ values), most of the GS2SLS models adhere quite well to the data. The fitness is not so good for weekdays late nighttime (0–6 h) and weekends daytime (6–20 h), but even in that case, the models capture a considerable part of the data.

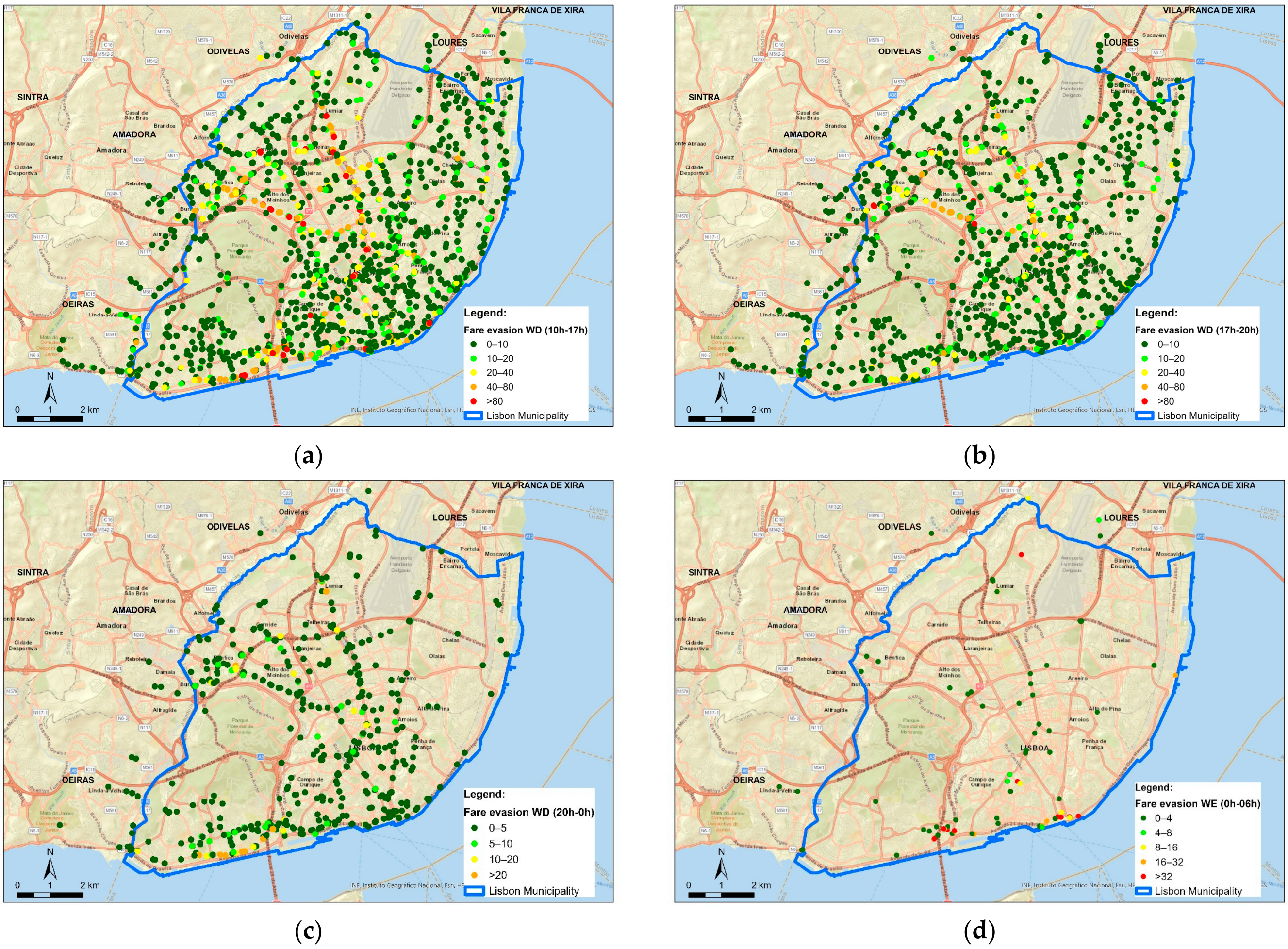
5. Discussion

Summing the above observations, firstly, it can be concluded that inspection actions are of major importance in assessing the state of fare evasion in the bus network of a mid-sized European capital city. Periodic inspection action and use of its results statistical models, such as the GS2SLS model this article presented, and persistent monitoring of the regression coefficients for this particular variable can be a valuable decision-aid tool for understanding whether the problem of fare evasion stabilized, decreased, or jumped out of control. An interesting managerial application of the GS2SLS model is it can be used to know whether the inspection actions saturated and may no longer be useful as deterrent, as suggested by [19,32]. If the model is run on new data with an increasing number of inspections and their regression coefficients do not decrease in comparison with previous ones, it is a hint that saturation may be happening and that other preventive or deterrent measures may be necessary.

Secondly, the period of the day and the distinction between weekdays and weekends influence how fare evasion is structured, in the sense that the statistical significance and trend direction of the other explanatory factors depend on period and type of day, although Tables 3 and 4 do not suggest any rationale besides the number of detected infringements and of inspection actions to link the various GS2SLS models.

Thirdly, little evidence was found of young age having an effect on the propensity to evade. This non-significance coupled to the fact that socio-economic data were not available due to data protection issues, which made it impossible to profile a typical premeditate evader.

Fourthly, concerning the spatial expression of fare evasion, the results hint at some dependence on the spatial component, not only through the significance of evasion detection and error autocorrelations in four of the models, but also because of reduced tendency for infringements on bus lines serving points of interest during rush hours. To understand whether spatial expressions of evasion can be identified, maps were derived, which are shown in Figure 5 below.



(**a**) (**b**) (**c**) (**d**)

Figure 5. Spatial distribution of detected evasion for periods with endogeneity. (**a**) Detected evasion, weekdays 10–17 h. (**b**) Detected evasion, weekdays 17–20 h. (**c**) Detected evasion, weekdays 20–0 h. (**d**) Detected evasion, weekends 0–6 h.

Recall that detected evasion (DE) exhibits spatial autocorrelation on the main term that is negative on the evening weekday rush hour (17–20 h) and positive on the weekend night period (0–6 h) and also positive spatial autocorrelation on the error term on two off-peak weekday periods (10–17 h, 20–0 h). The corresponding spatial patterns of DE are not straightforward to identify from the maps, but they appear a little more clearly in Figure 5a, which shows higher incidence of orange and red dots along two bus lines connecting the centre of Lisbon to the suburbs of Amadora and Odivelas (up to Lumiar) and along the Tejo River right bank. According to Figure 2, these bus lines are very frequently inspected, which brings up the following point. In a GS2SLS, when regressors have strong explanatory power and endogeneity plays a smaller role, it is possible that spatial patterns, albeit present, are hard to visualize or that those patterns are more tied to regressor influence than to endogeneity. The latter could be the case of Figure 5a because frequent inspections lead to more detection, which potentially explains the clustering along the above-mentioned bus lines. However, since this period has a statistically significant autocorrelation on the error term, undisclosed factors with a spatial expression must also exist, and therefore, the complete explanation of DE bears a spatial dependence that goes beyond the inspection/detection nexus. Possible factors with a spatial expression that might explain the error autocorrelation include socioeconomical variables, such as income level (which could not be included due to data protection).

In another period, Figure 5b, this one with anticorrelation on the DE main effect, a similar but less pronounced clustering is seen. In this case, the statistical significance does not fall on the error term but on the DE main effect itself, so, while not explicitly visible on the map, it nevertheless is there and suggests a deterrent effect of inspection actions because it is an anticorrelation. As to Figure 5c,d, spatial patterns of DE are even harder to visualize, and one must trust the statistical outcome, which is of a slight clustering tendency of DE in both cases. The final geographic factor is POI density, which is higher in downtown Lisbon, located to the southeast. Again, its effect (less DE) is not easily visible on the maps. Since the decrease effect applies at rush hours, it could be linked to commuting workers with valid bus passes, as POI-dense locations also tend to have higher job densities [42].

Summarizing on the spatial factors, the conclusion is that some spatial patterns exist that maintain statistical significance after controlling the number of inspection actions. However, it is not clear what those patterns are and, consequently, what might be their underlying cause (income levels are a possibility).

As a final note, Figure 6 shows a map of detected evasion.

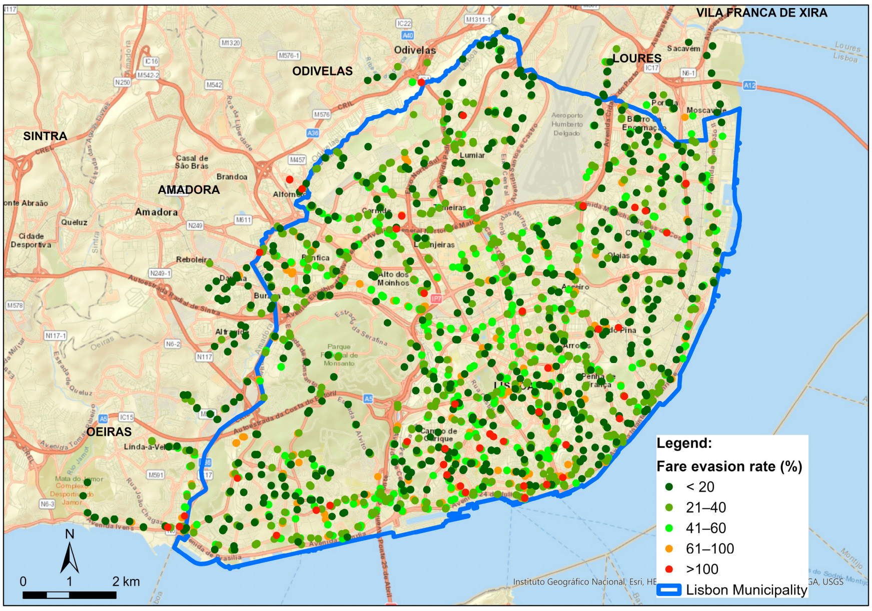


Figure 6. Global rate of fare evasion (DE/IA).

Figure 6 displays a scattered pattern which would suggest that the bus company is placing its inspectors correctly, carrying out more inspection actions on bus lines that indeed have a higher incidence of fare evasion.

Regardless, the results suggest policy implications of transferring inspectors from (POI-dense) downtown to north and northwest stops and intensifying inspections in the busier lines of weekdays 10–17 h period (due to the bus frequency factor), although these are mild suggestions as compared to monitoring the progression of the regression coefficients for the number of inspection actions and of the sign of Rho and Lambda W, whose negative value suggests effectiveness of the inspection actions.

6. Conclusions

Fare evasion has been addressed in several works, but previous studies were based on surveys limited in time, space, and size. The work proposed here was based on a larger database than most of the studies, whose size allowed us to analyze different periods and distinguish between working days and weekends and included the spatial component, usually a neglected factor. This made it possible to identify and relate patterns in the territory while controlling for autocorrelation by means of GS2SLS modelling. To the best of our knowledge, this is the first time such an approach has been developed.

Results showed that indeed the type of day (weekday or weekend) and period of the day lead to different models. The number of inspections actions had the highest impact on explaining the number of detected fare evasions, being the only variable that was consistently significant throughout all periods of the day and regardless of the type of day. It is expected that this variable remains significant in general, highlighting its importance as a decision-aid tool to monitor the state of evasion in the public bus lines. However, after controlling that variable, various statistically significant tendencies remained. For instance, it was also seen that bus frequency plays a role in off-peak hours and that, contrary to previous studies, young males do not seem to be more likely to infringe than the average citizen, suggesting that age is likely to be a local factor in evasion. However, the latter conclusion must be taken with a grain of salt, as data on age were obtained indirectly.

The influence of the spatial component, a key factor this research set out to assess, was also found to be statistically significant for several of the models. During the evening rush hour, the distribution of detected evasion exhibited spatial anticorrelation, suggesting a dissuasive effect of inspection actions. In other models, the endogenous effects turned up in the error term, hinting at spatial factors that are yet undisclosed. However, since the regressors have a high explanatory power as compared to endogeneity, spatial patterns of evasion were not visible from map representations, making it unclear what they might relate to geographically. The density of points of interest provides a tentative spatial explanation for those patterns, at least for the rush hour periods, which could be applicable to other case studies, but clearly more research must be carried out to fully understand the spatial aspects of fare evasion. This research showed that those aspects do play a statistically significant role, even in the presence of strong explanatory factors like number of inspections, and therefore warrant further study. A better knowledge of those patterns would also help understanding to what degree the conclusions of this article can be generalized in other locations.

Further putative factors of evasion with a spatial expression should certainly be tested out in follow-up research, datasets permitting. Another factor that may be interesting to consider is the direction of travel, as this might relate to evasion patterns of commuters. Deepening the analysis of whether inspection actions effectively deter fare evasion (and where) is another issue worth researching, as well as making an heterogenous analysis, i.e., deriving local regression coefficients. We hope to address these issues in the near future.

Author Contributions: Conceptualization S.F.; methodology S.F.; formal analysis S.F. and N.S.; investigation S.F. and N.S.; resources S.F.; data curation S.F.; writing—original draft preparation S.F. and N.S.; writing—review and editing S.F. and N.S.; visualization S.F.; supervision S.F. and N.S.; project administration S.F. and N.S. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Portuguese Foundation for Science and Technology under project grants nrs. UIDP/04427/2020 and UIDB/00308/2020. https://doi.org/10.54499/UIDP/04427/2020 (CITTA) and https://doi.org/10.54499/UIDB/00308/2020 (INESC-Coimbra). This research received no further external funding.

Data Availability Statement: Research data is not publicly available due to trade secrecy issues.

Acknowledgments: We would like to acknowledge CARRIS | Transporte Público de Lisboa for providing the data essential to develop this analysis.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

GS2SLS	Generalized spatial two-stage least-squares regression
IA	Inspection actions
DE	Detected evasion
POI	Point of interest
OSM	OpenStreetMap
INE	Instituto Nacional de Estatística—Portuguese statistics bureau

References

1. Killias, M.; Scheidegger, D.; Nordenson, P. The Effects of Increasing the Certainty of Punishment: A Field Experiment on Public Transportation. *Eur. J. Criminol.* **2009**, *6*, 387–400. [CrossRef]
2. Reddy, A.V.; Kuhls, J.; Lu, A. Measuring and Controlling Subway Fare Evasion: Improving Safety and Security at New York City Transit Authority. *Transp. Res. Rec.* **2011**, *2216*, 85–99. [CrossRef]
3. Celse, J.; Grolleau, G. Fare Evasion and Information Provision: What Information Should Be Provided to Reduce Fare-Evasion? *Transp. Policy* **2023**, *138*, 119–128. [CrossRef]
4. Abrate, G.; Fraquelli, G.; Meko, E.; Rodia, G. L'evasione Tariffaria Nel Trasporto Pubblico Locale: Un'analisi Empirica. In Proceedings of the Conferenza Società Italiana di Economia Pubblica, XX Riunione Scientifica, Pavia, Italy, 25–26 September 2008; pp. 25–26.
5. Suquet, J.-B. Drawing the Line: How Inspectors Enact Deviant Behaviors. *Emerald* **2010**, *24*, 468–475.
6. Currie, G.; Delbosc, A. An Empirical Model for the Psychology of Deliberate and Unintentional Fare Evasion. *Transp. Policy* **2017**, *54*, 21–29. [CrossRef]
7. Lee, J. Uncovering San Francisco, California, Muni's Proof-of-Payment Patterns to Help Reduce Fare Evasion. *Transp. Res. Rec.* **2011**, *2216*, 75–84. [CrossRef]
8. Munizaga, M.A.; Gschwender, A.; Gallegos, N. Fare Evasion Correction for Smartcard-Based Origin-Destination Matrices. *Transp. Res. A Policy Pract.* **2020**, *141*, 307–322. [CrossRef]
9. Egu, O.; Bonnel, P. Can We Estimate Accurately Fare Evasion without a Survey? Results from a Data Comparison Approach in Lyon Using Fare Collection Data, Fare Inspection Data and Counting Data. *Public Transp.* **2020**, *12*, 1–26. [CrossRef]
10. Cantillo, A.; Raveau, S.; Muñoz, J.C. Fare Evasion on Public Transport: Who, When, Where and How? *Transp. Res. A Policy Pract.* **2022**, *156*, 285–295. [CrossRef]
11. Almutairi, A.; Owais, M.; Ahmed, A.S. Notes on Bus User Assignment Problem Using Section Network Representation Method. *Appl. Sci.* **2024**, *14*, 3406. [CrossRef]
12. Guarda, P.; Galilea, P.; Paget-Seekins, L.; Ortúzar, J. de D. What Is behind Fare Evasion in Urban Bus Systems? An Econometric Approach. *Transp. Res. A Policy Pract.* **2016**, *84*, 55–71. [CrossRef]
13. Barabino, B.; Salis, S. Do Students, Workers, and Unemployed Passengers Respond Differently to the Intention to Evade Fares? An Empirical Research. *Transp. Res. Interdiscip. Perspect.* **2020**, *7*, 100215. [CrossRef]
14. Busco, C.; González, F.; Jaqueih, Y.; Jiménez, F.; Alonso, B. Understanding Transantiago Users' Motivations for Paying or Evading Payment of Bus Fares. *J. Public Transp.* **2022**, *24*, 100016. [CrossRef]
15. Delbosc, A.; Currie, G. Why Do People Fare Evade? A Global Shift in Fare Evasion Research. *Transp. Rev.* **2019**, *39*, 376–391. [CrossRef]
16. Barabino, B.; Salis, S.; Useli, B. Fare Evasion in Proof-of-Payment Transit Systems: Deriving the Optimum Inspection Level. *Transp. Res. B Methodol.* **2014**, *70*, 1–17. [CrossRef]
17. Delbosc, A.; Currie, G. Four Types of Fare Evasion: A Qualitative Study from Melbourne, Australia. *Transp. Res. Part F: Traffic Psychol. Behav.* **2016**, *43*, 254–264. [CrossRef]
18. Eddy, D. Fare Evasion Is It a Youth Issue? *Transit Aust.* **2010**, *65*, 1–7.
19. Cools, M.; Fabbro, Y.; Bellemans, T. Identification of the Determinants of Fare Evasion. *Case Stud. Transp. Policy* **2018**, *6*, 348–352. [CrossRef]
20. Perrotta, A.F. Transit Fare Affordability: Findings from a Qualitative Study. *Public Works Manag. Policy* **2017**, *22*, 226–252. [CrossRef]

21. Dai, Z.; Galeotti, F.; Villeval, M.C. Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation. *Manag. Sci.* **2018**, *64*, 1081–1100. [CrossRef]
22. Bucciol, A.; Landini, F.; Piovesan, M. Unethical Behavior in the Field: Demographic Characteristics and Beliefs of the Cheater. *J. Econ. Behav. Organ.* **2013**, *93*, 248–257. [CrossRef]
23. Barabino, B.; Salis, S.; Useli, B. What Are the Determinants in Making People Free Riders in Proof-of-Payment Transit Systems? Evidence from Italy. *Transp. Res. A Policy Pract.* **2015**, *80*, 184–196. [CrossRef]
24. Guzman, L.A.; Arellana, J.; Camargo, J.P. A Hybrid Discrete Choice Model to Understand the Effect of Public Policy on Fare Evasion Discouragement in Bogotá's Bus Rapid Transit. *Transp. Res. A Policy Pract.* **2021**, *151*, 140–153. [CrossRef]
25. Sasaki, Y. Optimal Choices of Fare Collection Systems for Public Transportations: Barrier versus Barrier-Free. *Transp. Res. B Methodol.* **2014**, *60*, 107–114. [CrossRef]
26. Milioti, C.; Panoutsopoulos, A.; Kepaptsoglou, K.; Tyrinopoulos, Y. Key Drivers of Fare Evasion in a Metro System: Evidence from Athens, Greece. *Case Stud. Transp. Policy* **2020**, *8*, 778–783. [CrossRef]
27. Clarke, R.V.; Contre, S.; Petrossian, G. Deterrence and Fare Evasion: Results of a Natural Experiment. *Secur. J.* **2010**, *23*, 5–17. [CrossRef]
28. González, F.; Busco, C.; Codocedo, K. Fare Evasion in Public Transport: Grouping Transantiago Users' Behavior. *Sustainability* **2019**, *11*, 6543. [CrossRef]
29. Bonfanti, G.; Wagenknecht, T. Human Factors Reduce Aggression and Fare Evasion. *Public Transp. Int.* **2010**, *59*, 28–32.
30. Fürst, E.W.M.; Herold, D.M. Fare Evasion and Ticket Forgery in Public Transport: Insights from Germany, Austria and Switzerland. *Societies* **2018**, *8*, 98. [CrossRef]
31. Larwin, T.F.; Koprowsky, Y. Off-Board Fare Payment Using Proof-of-Payment Verification. In *Sustaining the Metropolis: LRT and Streetcars for Super Cities, Proceedings of the 12th International Light Rail Conference, Salt Lake City, Utah, 11–13 November 2012*; Transportation Research Board 2013 Executive Committee Officers: Washington, DC, USA, 2013; pp. 71–83.
32. Porath, K.; Galilea, P. Temporal Analysis of Fare Evasion in Transantiago: A Socio-Political View. *Res. Transp. Econ.* **2020**, *83*, 100958. [CrossRef]
33. Freiria, S.; Sousa, N. The Impact of Accessibility Changes on Local Development: A Spatial Approach. *J. Transp. Geogr.* **2024**, *120*, 103975. [CrossRef]
34. Carris. *Relatório & Contas 2019*; Carris: Lisbon, Portugal, 2019.
35. Pourmonet, H.; Bassetto, S.; Trépanier, M. Vers La Maîtrise de l'évasion Tarifaire Dans Un Réseau de Transport Collectif. In Proceedings of the 11e Congrès International de Génie Industriel, Saint-Sauveur, QC, Canada, 26–28 October 2015.
36. Elhorst, J.P. Relever Le Niveau de l'économetrie Spatial Appliquée. *Spat. Econ. Anal.* **2010**, *5*, 9–28. [CrossRef]
37. Cordera, R.; Chiarazzo, V.; Ottomanelli, M.; dell'Olio, L.; Ibeas, A. The Impact of Undesirable Externalities on Residential Property Values: Spatial Regressive Models and an Empirical Study. *Transp. Policy* **2019**, *80*, 177–187. [CrossRef]
38. INE. *CENSOS 2021-Resultados Definitivos*; INE: Lisboa, Portugal, 2021.
39. Schneider, W. Bbbike Extract Service. Available online: https://extract.bbbike.org/ (accessed on 28 April 2025).
40. Hensher, D.A.; Li, Z.; Mulley, C. Drivers of Bus Rapid Transit Systems—Influences on Patronage and Service Frequency. *Res. Transp. Econ.* **2014**, *48*, 159–165. [CrossRef]
41. Owais, M.; Osman, M.K. Complete Hierarchical Multi-Objective Genetic Algorithm for Transit Network Design Problem. *Expert Syst. Appl.* **2018**, *114*, 143–154. [CrossRef]
42. Wang, Z.; Zheng, J.; Han, C.; Lu, B.; Yu, D.; Yang, J.; Han, L. Exploring the Potential of OpenStreetMap Data in Regional Economic Development Evaluation Modeling. *Remote Sens.* **2024**, *16*, 239. [CrossRef]