



Fare evasion strategies, perceived safety, and travel choices in public transport: evidence from a vignette experiment

Benedetto Barabino^a, Martina Carra^a , Jérémy Celse^{b,*} , Roberto Ventura^a

^a DICATAM, Università di Brescia, Italy

^b ESSCA School of Management, France

ARTICLE INFO

JEL:

R41

K42

R48

C91

Keywords:

Fare evasion

Public transport compliance

Perceived safety and security

Deterrence strategies

Ride-sharing competition

ABSTRACT

Fare evasion poses persistent challenges for public transport systems, with implications for financial sustainability, social norms, corporate image, and perceived safety and security. While deterrence strategies such as increasing inspection frequency or fine severity are widely used, their effects on passengers' perceptions and behavioural spillovers remain weakly evidenced. This study examines how anti-fare evasion cues influence travel preferences, perceived safety and security, and fare-evasion intentions, accounting for time of day and competition from a ride-sharing service (RSS). Four research questions were examined through a vignette-based randomised experiment. Participants chose among a non-manipulated bus route, a manipulated route (enforcement or service cue), and an RSS alternative across seven conditions crossing enforcement cues (inspection, fines, neutral) and period (day/night). The results showed that the effects of anti-fare evasion cues were small and statistically uncertain. Neither inspections nor fines produced clear improvements in perceived safety or security (RQ1), and the enforcement type showed no consistent direct effect on the choice of the manipulated route (RQ2). Night-time conditions did not meaningfully change safety or risk perceptions, nor did they reliably moderate the association between perceived route safety and route choice (RQ3). Fare evasion was more frequent at night and less frequent on the manipulated route yet contrasts between Fine and Inspection treatments remained small and imprecisely estimated (RQ4). The RSS garnered only 3.73% choices amid price differentials. Thus, deterrence failed to bolster reassurance or ridership proxies, with choices driven by safety and security perceptions independently of cues. Night primarily led to increased evasion without corresponding gains in safety and security. Policies should pair targeted enforcement with service enhancements, particularly nocturnal, in multi-modal contexts.

1. Background

Fare evasion poses a significant challenge for Transit Authorities (TAs) and Public Transportation Companies (PTCs) worldwide, with implications for revenue losses, operational efficiency, corporate image, social equity, and safety and security. It can generate substantial financial deficits, undermine the long-term sustainability of public transport systems and potentially lead to fare increases or service reductions, which may negatively affect passenger satisfaction and, in some circumstances, contribute to further fare evasion

* Corresponding author.

E-mail addresses: benedetto.barabino@unibs.it (B. Barabino), martina.carra@unibs.it (M. Carra), jeremy.celse@essca.fr (J. Celse), roberto.ventura@unibs.it (R. Ventura).

<https://doi.org/10.1016/j.tra.2026.105244>

Received 4 February 2026; Received in revised form 25 June 2026; Accepted 1 September 2026

0965-8564/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

List of acronyms

AME	Average Marginal Effect
ARM	Experimental Arm (Inspection of Fine conditions)
BSL	Baseline condition
CI	Confidence Interval
CFA	Confirmatory Factor Analysis
GSEM	Generalised Structural Equation Models
ITT	Intention-To-Treat
OLS	Ordinary Least Squares
OR	Odds Ratio
POP	Proof-Of-Payment
PP	Per-Protocol
PT	Public Transport
PTC	Public Transport Company
RQ	Research Question
RSS	Ride Sharing Service
SaRI	Safety Risk Index
SeRI	Security Risk Index
TA	Transit Authority
TOST	Two One-Sided Tests

List of variables

INSP_D	inspection condition (day)
INSP_N	inspection condition (night)
FINE_D	fine condition (day)
FINE_N	fine condition (night)
IRR_D	irrelevant condition (day)
IRR_N	irrelevant condition (night)
Safety_LA	perceived safety of the manipulated route
Safety_LB	perceived safety of standard (or not-manipulated) route
Safety_RSS	perceived safety of ride-sharing service
Δ Safety	difference in perceived safety between manipulated and standard route
Δ Safety_ITT	ITT safety difference
Δ ProbInspect	difference in perceived inspection probability between the manipulated and standard route

(Allen et al., 2019; Clarke et al., 2010; Guzman et al., 2021). To address the issue, TAs/PTCs have traditionally relied on “direct” deterrence and enforcement-based strategies, primarily through ticket inspections and the imposition of fines (or penalties) on offenders. These measures are grounded in rational choice theory, which posits that individuals weigh the expected costs and benefits of non-compliance (Becker, 1968). Accordingly, increasing the likelihood of detection, perceived detection, or charging higher than usual fines are considered effective levers to discourage fare evasion (Beyleveld, 1979a, 1979b; Bijleveld, 2007; Freiria and Sousa, 2025; Hauber, 1993).

Several studies have examined the effectiveness of these deterrence-based strategies, primarily designed to reduce fare evasion by increasing the expected cost of non-compliance, i.e., to reduce fare evasion by targeting those who evade fares or those who could evade; however, they yielded both convergent and divergent results (Barabino et al., 2024a). Beyond their direct deterrent effects, enforcement measures may also influence passengers’ perceptions of safety and security (on board and at stops) and, more broadly, their travel experience, and mobility choices. Indeed, Clarke et al. (2010) and Porath and Galilea (2020) highlighted that fare evasion is not merely an economic issue; it also contributes to an environment perceived as less safe and secure, affecting perceived safety, security, and service quality (e.g., Reddy et al., 2011). Freiria and Sousa (2026) found that inspection actions could improve both detection effectiveness and perceived security. However, these benefits vary across locations and operational settings, indicating a strong context-dependent effect. Moreover, fare evasion reinforces feelings of disorder, insecurity, and unsafety among passengers, thereby influencing ridership decisions. It can erode trust in public institutions and contribute to a sense of disorder within transit environments (Hirsch et al., 1999; van Dijk and Junger-Tas, 1988). According to Johnson and Patterson (2022), the perceived presence of disorder, crime, or deviant behaviours (as fare evasion can be) within public transport systems often conflicts with expectations of

safety and security.¹ This is particularly among middle-class users, who tend to be more sensitive to signs of disorder in public spaces. Therefore, fare evasion (and the strategies adopted to address it) may also affect passengers who comply with fare rules. It suggests that enforcement policies cannot be assessed solely in terms of direct compliance outcomes, as they can also interact with passengers' emotional and cognitive responses to safety and order, producing "indirect" economic effects.

Previous studies have shown that perceptions of safety and security are important determinants of public transport use (Loukaitou-Sideris, 2014; Newton, 2004). Nevertheless, limited attention has been devoted to understanding how fare evasion and fare enforcement strategies shape these perceptions and, in turn, influence passengers' mobility choices.

Conceptually, fare-evasion enforcement may influence behaviour through both deterrent and perceptual (or guardianship) pathways, each targeting different actors and cognitive processes within the transport environment (see [Supplementary Material – Appendix A](#)). From a deterrence perspective, inspections and fines increase the expected cost of non-compliance for potential fare evaders by raising either the perceived probability of detection or the severity of punishment (Becker, 1968; Polinsky and Shavell, 2000). Beyond this compliance mechanism, enforcement cues may also affect observing, fare-compliant passengers by signalling the degree of monitoring and behavioural regulation in the transport environment. Building on routine activity theory, and consistent with transport research extending this logic to perceived security, visible inspections may act as guardianship cues that reduce perceived disorder and vulnerability (Ceccato et al., 2022; Cohen and Felson, 1979; Newton, 2004). Disorder-based accounts of perceived insecurity offer a complementary logic: visible signs of unchecked rule-breaking may signal weak informal control and undermine perceived security, whereas active enforcement may indicate that deviance is not tolerated (Skogan, 1992; Wilson and Kelling, 1982). This dual-pathway framework provides the theoretical rationale for why inspection- and fine-based strategies may yield divergent perceptual consequences for observing passengers: inspections provide a concrete, visible monitored presence that feeds directly into the guardianship mechanism; fines operate more abstractly through regulatory severity and perceived sanction costs for would-be evaders, without necessarily generating an immediate, physical sense of environmental control for the general travelling public. These mechanisms may be particularly salient at night, when low-light conditions structurally alter cue salience, environmental vulnerability, and the subsequent perceptual response of passengers, making vulnerability and disorder cues likely to weigh more heavily on travel evaluations. Therefore, we treat perceived security and, more cautiously, perceived safety as potential mediators among enforcement cues, route choice, and fare-evasion intentions, while recognising that enforcement may also backfire if interpreted as evidence of conflict or mistrust (Fürst and Herold, 2018; Suquet, 2010).

Consequently, two interrelated topics emerge: (i) the links among fare evasion, fare enforcement strategies, and perceived safety and security within public transport systems; and (ii) how behavioural consequences of these perceptions may influence the modal choice of users and the relative competitiveness of public transport with respect to alternative mobility options, such as ride-sharing services (RSS).

Regarding topic (i), the literature on fare evasion control and perceptions of safety and security is limited (Barabino et al., 2024b; Ceccato et al., 2022; Ceccato and Loukaitou-Sideris, 2020; Loukaitou-Sideris, 2014; Newton, 2004; Smith and Cornish, 2006). Only a handful of studies face security issues related to fare evasion. For instance, Barabino et al. (2020) framed it as the protection of public transport staff (particularly ticket inspectors and bus drivers) from physical or verbal assaults by fare evaders or uncooperative passengers. Such incidents, which undermine staff authority and disrupt enforcement activities, are influenced by contextual factors like time of day, location, and passenger density. These findings are consistent with studies showing that perceived security in public transport declines significantly during night-time travel due to reduced visibility, lower passenger density, and heightened perceptions of vulnerability to crime or harassment (Ceccato et al., 2022; Ceccato and Loukaitou-Sideris, 2020; Killias et al., 2009; Newton, 2004). Moreover, security concerns are particularly relevant for certain demographic groups, such as women and young adults, who report greater reluctance to use public transport after dark (Loukaitou-Sideris, 2014; Smith and Cornish, 2006). These patterns are consistent with behavioural research that highlighted how moral and risk-related conducts (including fare evasion) are sensitive to environmental cues, e.g., darkness and perceived social control, which tend to vary across time and space (Keuschnigg and Wolbring, 2015; Zhong et al., 2010). Further empirical evidence highlighted the close connection between fare enforcement and on-board security. A survey reported that about 67% of assaults on bus operators are related to fare enforcement (Nakanishi and Fleming, 2011). Similarly, Macé (1997) linked insecurity on buses to tensions between lower-income passengers and drivers. Authors such as Bonfanti and Wagenknecht (2010); Mauri et al. (1984); Nakanishi and Fleming (2011) suggested that assaults and fare-related conflicts can be mitigated through technological (e.g., cameras, alarms, protective cabins) and organisational (e.g., police cooperation, visible security presence, staff training) measures. These strategies are expected to enhance both staff security and fare compliance by reinforcing perceived control and authority on board (Ceccato and Loukaitou-Sideris, 2020). By contrast, safety-related events (e.g., bus crashes) can undermine passengers' confidence in bus services. Since safety is also a key dimension of service quality, offering trips that are perceived as safer than those by car can positively influence users' perceptions and service usage (Barabino et al., 2024b; Barabino et al., 2020).

Regarding topic (ii), perceived safety and security may play a role in influencing modal choice and the attractiveness and competitiveness of public transport relative to alternative mobility options such as RSS. Services such as Uber or Lyft have transformed urban mobility by offering flexible, on-demand alternatives that often prioritise perceived safety, security, and cost convenience (Clewlow and Mishra, 2017; Shaheen and Cohen, 2019). Empirical evidence suggests that safety and security concerns are a key driver

¹ Safety refers to system property and management practices aimed at preventing accidents and crashes and minimising their consequences; Security refers to the systemic prevention, deterrence, detection and responses to criminal acts targeting passengers, staff, vehicles and infrastructures to reduce the incidence, fear, and harm of crime across the network to an acceptable risk level (Eck & Clarke, 2019).

of RSS adoption, particularly during night-time travel, as passengers perceive these services as offering greater personal safety and security than public transport (Alemei et al., 2018; Dias et al., 2017). Including RSS in the choice set enables us to capture realistic substitution patterns and assess whether anti-fare evasion strategies can enhance the competitiveness of public transport relative to these alternatives. Comparing buses and RSS is crucial for understanding how enforcement measures interact with perceptions of safety and security, thereby shaping modal preferences within an increasingly multimodal transport ecosystem.

1.1. Literature gaps

All the previous literature has presented fare evasion control mainly through deterrence- and enforcement-based approaches grounded in rational choice theory, while a parallel body of research has emphasised the role of perceived security in shaping public transport use.

However, some gaps persist.

First, the primary strategies employed, ticket inspections and fines, are well documented in terms of deterrence from the viewpoint of TAs/PTCs, yet their acceptability and perceived desirability from the passenger perspective remain underexplored. Do passengers prefer PTCs that prioritise inspections over fines? To what extent do they expect TAs/PTCs to face fare evasion, and can these strategies serve as competitive advantages against emerging mobility providers such as RSS?

Second, only a handful of studies have examined whether enforcement strategies influence passengers' security perceptions or interact with them to shape fare compliance and modal choice; most empirical work instead estimates deterrence effects on evasion rates and detection rather than on perceived safety and security. For example, a quasi-experiment in Münster found that raising inspection levels reduced evasion, that afternoon drops in inspection raised evasion, and that switching from official uniforms to civilian clothing increased detections for "no valid ticket" while decreasing "ticket forgotten," yet none of these outcomes included riders' perceived safety or security measures (Keuchel and Swertz, 2020). Qualitative work in Brussels documents how anti-evasion controls, surveillance, and online alerting communities shape experiences of control, stigma, and fear-based exclusion, indicating that security perceptions and access are likely intertwined (Assaf and Van den Broeck, 2022). Moreover, to our knowledge, no empirical study has directly linked the prevalence of fare evasion to perceived safety issues. Understanding the relationship between fare evasion and safety and security perceptions is particularly relevant: observing rule-breaking and the manner of its control may signal (a lack of) order and system care, potentially affecting riders' sense of safety and security and their use of the system, a hypothesis consistent with evidence on control salience and with discussions of fear-based exclusion on public transport.

1.2. Goal

This study addresses these gaps by analysing passengers' preferences for alternative anti-fare evasion strategies in public transport. To be more specific, this study aims to answer the following research questions (RQs):

RQ1: Do anti-fare evasion strategies increase feelings of safety and security?

Enforcement-oriented messages (inspection and fine cues) are expected to increase perceived safety and security relative to a neutral baseline. Namely, we conjecture that participants report higher levels of safety and security when PTCs implement anti-fare evasion measures (either by increasing fines or by increasing ticket inspection operations).

RQ2: Do anti-fare evasion enforcement strategies indirectly affect public transport choices through changes in perceived safety and security?

The relationship between perceived safety/security and public transport use is generally well established (Currie et al., 2013; Delbosc and Currie, 2012). However, this study treats it as a behavioural mechanism through which anti-fare evasion enforcement strategies may indirectly influence travel choices. Higher perceived safety and security are expected to increase the likelihood of choosing public transport, particularly the bus route associated with stronger enforcement. Namely, because of anti-fare evasion strategies implemented by the PTC (either by increasing fines or ticket inspection operations), we expect participants to report higher levels of safety and security, which push them to travel through the manipulated route.

RQ3: Are the effects of feelings of safety and security more important during night travels?

The effects of enforcement strategies on safety/security perceptions and, through them, on transport choices are expected to be stronger at night than during the day. Namely, we hypothesise that anti-fare evasion strategies are more likely to increase safety and security feelings during night travels than during day travels.

RQ4: Do Inspection and Fine enforcement reduce fare evasion?

Enforcement strategies are expected to reduce fare evasion by increasing the expected chance of being caught evader. Namely, we assume that participants are more likely to fare evade when PTCs do not engage in anti-fare evasion strategies.

By manipulating enforcement cues (inspection, fines, neutral) and time of day (daytime vs night-time), this study examines how these strategies affect perceived safety and security and their potential behavioural implications, thereby enhancing the ecological validity of the experimental design. These effects are empirically examined using a dataset of 300 questionnaires collected from business school students in France. Notably, the findings should be interpreted within this defined socioeconomic context rather than as representative of the overall population of public transport users, as fare evasion behaviour is strongly influenced by socioeconomic conditions and may vary significantly across developed and developing contexts (e.g., Barabino and Salis, 2023; Delbosc and Currie, 2016; Munizaga et al., 2020).

1.3. Implications for theory and practice

This study aims to advance understanding of fare evasion by linking deterrence theory to behavioural insights into perceived safety and security.

Specifically, from a theoretical perspective, this study addresses a research gap that, to the authors' knowledge, has not yet been thoroughly examined. In addition, it enhances the understanding of some factors influencing passengers' safety and security perceptions to evade fares. For instance, modelling safety and security perceptions along transit routes could refine the formulation of optimisation models dealing with the current issue in patrol inspection strategies, both in massive and collective PT (Escalona et al., 2024; Escalona et al., 2025).

From a practical standpoint, evaluating safety and security perceptions can assist PTCs in tailoring specific countermeasures against potential fare evaders. For example, inspection-based measures could foster greater trust and compliance than fine-based deterrence (if any), highlighting the need for visible, fair, and integrated control strategies. Combining enforcement with environmental and operational improvements, such as lighting, staff presence, and reduced waiting times, can simultaneously enhance perceived safety and fare compliance, strengthening overall system legitimacy and ridership. Moreover, exploring safety and security perceptions in transport choices represents a critical lever for influencing mobility behaviour. RSS companies such as Uber and Lyft have recognised this by introducing women-only services during night hours, addressing gender-specific concerns about personal safety. TAs/PTCs must similarly acknowledge and respond to the diverse motivations and perceptions shaping passenger decisions. For instance, if men and women react uniformly to security-related stimuli, it often leads to ineffective strategies, as empirical evidence frequently invalidates such assumptions. Therefore, a nuanced understanding of these perceptual differences is essential for designing inclusive, trust-enhancing interventions that improve both perceived safety and overall system legitimacy.

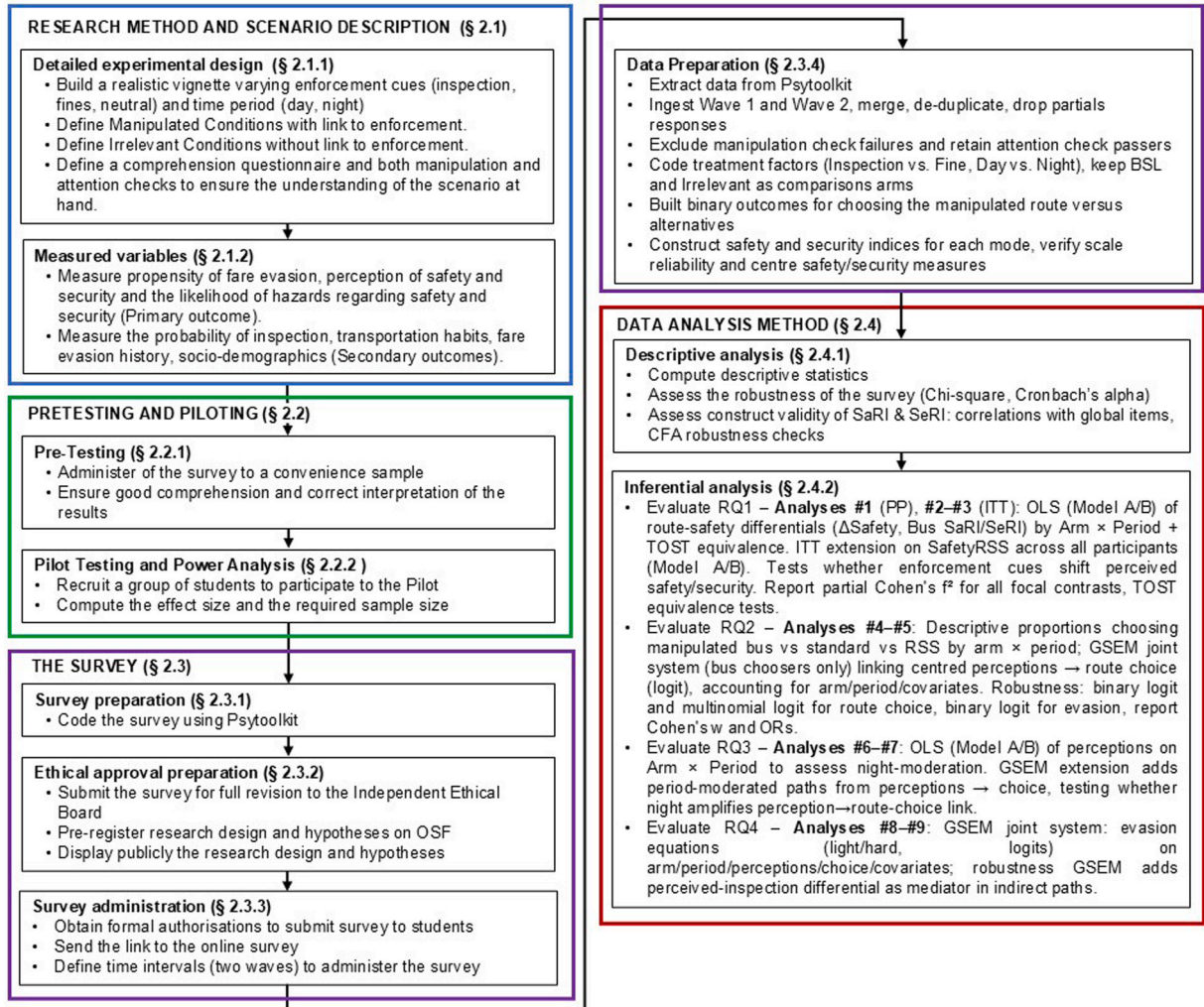


Fig. 1. Flowchart of the proposed methodology.

1.4. Organisation

The remainder of the paper follows. [Section 2](#) presents the research methodology. [Section 3](#) reports the results. Finally, [Section 4](#) discusses and concludes this study.

2. Methodology and data

This section outlines the methodology applied to address the research questions. It first introduces the research method, then describes the experimental design, pre-test, and piloting. Finally, it presents the survey method and the data analysis procedures. The overall methodology is summarised in [Fig. 1](#) and detailed in the subsequent sections.

2.1. Research method and scenario description

This study employs a vignette-based randomised experiment to investigate how perceived safety and security, and fare evasion intentions, are influenced by anti-fraud interventions (increased penalties or inspection frequency) across day and night periods in public transport settings.

A vignette-based randomised experiment is well-suited for capturing real-world complexity, enabling standardised scenario exposure while controlling for extraneous variables ([Atzmüller and Steiner, 2010](#)). This experiment was selected for its ability to present participants with realistic, controlled scenarios that closely mirror real-world decision-making contexts. This approach enables precise manipulation of key variables while maintaining ecological validity, facilitating the examination of causal effects on behaviour ([Finch, 1987; Hughes and Huby, 2004; Steiner et al., 2017](#)). Additionally, vignettes provide a less direct and more anonymous way to explore sensitive topics such as fare evasion, thereby reducing social desirability bias and enhancing the reliability of self-reported intentions. Recent studies are referring to vignette studies to explore transportation decisions ([Celse and Grolleau, 2023](#)).

In this experiment, participants are presented with a realistic scenario in which they have recently relocated to a new city for academic purposes. After a meal, they must choose a mode of transport to return home from a restaurant, either during the day or late at night (depending on the condition). Participants receive detailed information about two proximate bus routes and an alternative RSS. To ensure comparability and interpretability of outcomes, all transport options are designed to arrive home at approximately the same time. However, while total travel time is held approximately constant across alternatives, the composition of travel time differs (e.g., waiting time versus in-vehicle time), which may influence perceived generalised cost, as walking and waiting time are typically valued more negatively than in-vehicle time ([Ceder, 2019; Ortuzar and Willumsen, 2011](#)). However, such differences are considered part of the scenario realism rather than experimental manipulations. This configuration is identical across all experimental conditions and not varied across treatments (i.e., Inspection, Fine, Irrelevant, Baseline), so any resulting differences in perceived generalised cost remain constant and do not affect the identification of treatment effects.

Importantly, although bus ticket prices for both routes are identical, using the RSS is approximately five times more expensive, reflecting prevailing local economic conditions. This cost differential encourages participants to consider multiple attributes simultaneously (time, cost, enforcement, and perceptions), rather than relying on a single dimension. This scenario compels participants to make decisions under uncertainty, balancing considerations of the likelihood of fare enforcement and perceptions of personal safety and security, thereby simulating typical urban transport decision-making contexts ([Berryessa and Caplan, 2020; Louviere et al., 2000](#)). A concrete example of an actual vignette presented to respondents is provided in the [Supplementary Material – Appendix B](#).

2.1.1. Detailed experimental design

The experiment manipulates two factors within the treatment arms: message content (Inspection, Fine or Irrelevant) and time of day (Day or Night). This results in a 3×2 factorial design, complemented by an additional baseline control condition that does not interact with any factor, for a total of seven conditions. The experiment adopts a between-subjects design, in which each participant is

Table 1

Overview of the experimental conditions.

Condition (abbreviation)	Manipulation	Inspection Frequency	Fine Level	Irrelevant Attribute	Time of Day
Baseline (BSL)	No information	Same for both routes	Same for both routes	None	Unspecified
Inspection Day (INSP_D)	More inspections on a route	Inspection is twice higher in the manipulated route	Same for both routes	None	Day
Inspection Night (INSP_N)	More inspections on a route	Inspection is twice higher in the manipulated route	Same for both routes	None	Night
Fine Day (FINE_D)	Fines are higher on a route	Same for both routes	Fines are twice higher in the manipulated route	None	Day
Fine Night (FINE_N)	Fines are higher on a route	Same for both routes	Fines are twice higher in the manipulated route	None	Night
Irrelevant Info Day (IRR_D)	Free Wi-Fi service on a route	Same for both routes	Same for both routes	Free Wi-Fi service on the manipulated route	Day
Irrelevant Info Night (IRR_N)	Free Wi-Fi service on a route	Same for both routes	Same for both routes	Free Wi-Fi service on the manipulated route	Night

randomly assigned to one of seven conditions that reflect different combinations of fare control, including a manipulated route and a non-manipulated (or standard) route. The manipulated route is defined as the route to which the experimental treatment is applied. Depending on the assigned condition, this route is associated with increased inspection intensity (INSP_D and INSP_N), higher fines (FINE_D and FINE_N) or a neutral attribute (IRR_D and IRR_N), while the standard route remains unchanged and serves as a reference.

Experimental manipulations vary the type of information provided to participants. More precisely, under the *Inspection* conditions, the manipulated route is described as having twice as many ticket inspections as the other. In *Fine* conditions, the manipulated route features fines that are twice as high. In the *Irrelevant* conditions, the manipulated route simply offers free Wi-Fi, a cue unrelated to fare enforcement. The Baseline (BSL) condition provides no additional information about routes, inspections, or services, thereby serving as a neutral reference point. Table 1, which is self-explanatory, summarises all conditions.

More in detail, the BSL and Irrelevant conditions jointly strengthen the validity of the experiment. The BSL condition offers a zero-information counterfactual against which all other cues (inspection, fines, or any neutral cue) can be benchmarked. It serves two main functions that strengthen internal validity. First, it provides an absolute reference level for choices (and perceived safety and security) in the absence of manipulation, thereby improving the interpretability of effect sizes and facilitating transparent robustness checks. Second, it disentangles content effects from mere salience/demand effects generated by providing additional information. To control for the potential influence of providing additional, but non-decision-relevant information, the experiment includes two “irrelevant information” conditions. In these conditions, participants received a neutral informational cue (the availability of free Wi-Fi on one of the bus routes) that bore no direct relation to fare evasion risks or enforcement measures. The purpose of these conditions (IRR_D and IRR_N) is to isolate and test whether any behavioural changes observed in the experimental treatments could be attributed to the presence of additional salient information per se, rather than the content of that information: if behaviour under Irrelevant are statistically similar to those in BSL, observed differences in Inspection and Fine conditions are credibly attributable to deterrence content rather than to “more text” or experimenter demand. This approach strengthens internal validity by ruling out confounds linked to the mere presence of extra information rather than its substantive relevance to travellers’ decisions regarding ticket purchase and compliance.

Day and night situations are included to account for literature indicating that safety and security, and moral behaviour, can be modulated by time, perceived ambient risk, and environmental cues (e.g., darkness, social threat) (Bonfanti and Wagenknecht, 2010; Killias et al., 2009). Randomisation ensures group equivalence.

On the same route, every information provided was randomised, including the route numbers and the order of choices offered to participants, to prevent anchoring biases.

Finally, we carefully incorporated elements to enhance data quality. First, we included a comprehension questionnaire (4 multiple-choice questions) after presenting the situation and before participants were invited to decide, to ensure they correctly understood the situation as well as to emphasise the manipulation. In line with best-practice recommendations, we added attention and manipulation check questions to ensure participants were attentive while completing the survey and noticed the manipulation (Aguinis and Bradley, 2014; Hughes and Huby, 2004). As recommended in standard processes, data from participants failing at these two questions will be discarded from the following analyses.

2.1.2. Measured variables

The experiment measures several constructs using single-items or psychometric scales consistent with the behavioural economics and moral psychology literatures. Specifically, to fulfil our research objectives, we measure both fare evasion propensity and perceived safety and security as primary outcomes. Moreover, we included some moderators and variables as controls by asking participants to answer questions unrelated to their transportation choice at the end of the study, as secondary outcomes. Both are described in what follows.

As for the primary outcomes, fare evasion propensity is measured by asking participants, if they choose to return home by bus, to indicate their ticket purchase and ticket validation behaviour, thereby directly indicating willingness to evade fare under the assigned scenario. Indeed, especially in Proof-of-Payment (POP) systems, travellers can evade fares by travelling without purchasing a ticket, without validating it, or by travelling with an inappropriate fare (tariffs depend on economic conditions). We chose to focus on the first two forms of fare evasion and to measure the two behavioural intentions, because these forms are the most widespread among bus passengers (Alade et al., 2020; Barabino and Salis, 2023; González et al., 2019; Sterner and Sheng, 2013).

Safety and security were operationalised as two related but distinct constructs, each measured for the bus options and the RSS using brief 7-point Likert-type items and a global “feel safe/secure” rating for each option. This approach aligns with transit scholarship that treats safety/security as a multidimensional perception spanning the entire trip, sensitive to setting and time of day, and consequential for user behaviour.

Safety was measured as the perceived likelihood of accidental or operational harm captured by six items per mode (Currie et al., 2013; Sundling et al., 2024): 0 – poor maintenance, 1- defective equipment, 2- Injured on board, 3- boarding/alighting injuries, 4 – road crashes, and 5 – serious injuries in a crash. Each item was evaluated with a 1–7 response scale, where higher values indicate a greater perceived likelihood of the unsafe event.

Security was measured as the perceived likelihood of crime and disorder events, captured by five items per mode (Elvik, 2019; Fruin et al., 1994): 0 – verbal aggression, 1 – physical aggression, 2 – theft, 3 – witnessing violent altercations, and 4 – incidents involving intoxicated passengers. Each Item was evaluated with a 1–7 response scale, again coded so higher values reflect a greater perceived likelihood of the insecure event.

For each option (i.e., bus or RSS), two composite indices were formed by averaging items: an operational Safety Risk Index – SaRI (accident/operational items) and a Security Risk Index- SeRI (crime/disorder items), with higher scores indicating greater perceived

unsafe/unsecure risk.

Therefore, SaRI and SeRI were treated as manifest composite indices: they are deterministic transformations of their respective item responses, not freely estimated latent variables. The item sets were designed to capture two theoretically distinct yet related domains (operational safety risk and interpersonal security risk) and the averaged scores preserve the original 1–7 response metric while providing directly reproducible proxies for these constructs. The use of averaged composite indices rather than latent factor scores was motivated by considerations of reliability, parsimony, and transparency. First, summated scales are commonly used when items are intended to represent the same conceptual domain and show acceptable internal consistency, because unit-weighted averages provide transparent and reproducible proxies for the underlying constructs (Cronbach, 1951; Nunnally and Bernstein, 1994; Raykov and Marcoulides, 2011). Accordingly, internal consistency was assessed for each multi-item scale. Second, the structural architecture of the research design calls for an empirical framework capable of jointly accommodating continuous perception outcomes and binary behavioural outcomes within a single estimated system. Estimating each equation separately (e.g., via individual linear models for perception outcomes and separate logit models for behavioural outcomes) would still permit single-equation inference, but it would not permit inference about indirect effects within a unified likelihood framework. Recovering indirect effects would then require auxiliary procedures external to the model, such as delta-method approximations or bootstrapping of products of coefficients (MacKinnon, 2012; Preacher and Hayes, 2008). Conventional covariance-based SEM accommodates categorical variables through threshold models in which observed binary responses are treated as indicators of underlying continuous latent variables (Muthén, 1984). That approach is appropriate when binary responses serve as indicators of a latent construct. In the present context, however, route choice and fare-evasion decisions are not indicators of an underlying latent disposition but the substantive behavioural outcomes of interest. A Generalised SEM (GSEM) with logit links is therefore more appropriate because it models these binary outcomes directly as dependent variables within the structural system, while allowing the perception equations to be estimated as approximately continuous outcomes (Rabe-Hesketh and Skrondal, 2012). Perception measures, although bounded on a 1–7 response scale, were treated as approximately continuous given the number of response categories, the composite construction of the indices, and standard practice in applied behavioural research (Norman, 2010; Sullivan and Artino, 2013). More specifically, the GSEM jointly estimates four perception equations (Δ Safety, SaRI, SeRI, and Safety_RSS) and three behavioural equations (route choice, light fare evasion, and hard fare evasion), with enforcement arm, travel period, and individual-level covariates entering as upstream predictors across the system. Given this multi-equation specification, incorporating latent measurement sub-models for SaRI and SeRI would have substantially increased the number of free parameters, raising risks of identification difficulties, convergence instability, and imprecise estimation in the non-linear behavioural equations. For this reason, SaRI and SeRI were treated as composite indices in the main specification. Treating SaRI and SeRI as observed composites therefore preserved estimation feasibility while retaining the simultaneous multi-equation structure and the mixed-outcome capability that motivate the GSEM specification, and while keeping the

Table 2

Description of items used to measure safety and security perceptions.

Construct	Items	Item Number	Score
Summary Indicator + Convergent validity	Would you feel safe/secure (i.e. not in danger or not in a risky situation) travelling on [Route A/ Route B/ using a RSS] to go home?	Safety_LA (Manipulated) Safety_LB (Standard) Safety_RSS	1 (Absolutely not) ... 7 (Absolutely yes)
Safety	Traveling on a [bus or RSS] that is poorly maintained.	Transp_Safety_0_Bus Transp_Safety_0_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Safety	Traveling on a [bus or RSS] with faulty equipment (e.g., air conditioning, seat armrests, seat belts...).	Transp_Safety_1_Bus Transp_Safety_1_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Safety	Being physically injured by [bus or RSS] equipment (e.g., getting your hand stuck in an armrest, bumping into a pole...).	Transp_Safety_2_Bus Transp_Safety_2_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Safety	Being physically injured while getting on or off a [bus or RSS].	Transp_Safety_3_Bus Transp_Safety_3_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Safety	Having a road crash when traveling by [bus or RSS].	Transp_Safety_4_Bus Transp_Safety_4_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Safety	Being seriously injured in the event of a road crash when traveling by [bus or RSS].	Transp_Safety_5_Bus Transp_Safety_5_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Security	Being verbally assaulted when traveling by [bus or RSS].	Transp_Safety_6_Bus Transp_Safety_6_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Security	Being physically assaulted when traveling by [bus or RSS].	Transp_Safety_7_Bus Transp_Safety_7_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Security	Having an item stolen when traveling by [bus or RSS].	Transp_Safety_8_Bus Transp_Safety_8_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Security	Witnessing a violent altercation between passengers (verbal or physical assault) when traveling by [bus or RSS].	Transp_Safety_9_Bus Transp_Safety_9_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)
Security	Witnessing incidents involving drunk passengers when traveling by [bus or RSS].	Transp_Safety_10_Bus Transp_Safety_10_RSS	1 (Extremely unlikely) ... 7 (Extremely likely)

Note: SaRI is computed as the arithmetic mean of Safety items (items 0–5) for each transport mode (Bus, RSS); SeRI is computed as the arithmetic mean of Security items (items 6–10) for each transport mode. Both indices are deterministic composites of observed item responses, not latent factor scores. The global single-item ratings (Safety_LA, Safety_LB, Safety_RSS) “Would you feel safe/secure travelling on...?” are standalone holistic outcomes and serve as convergent-validity anchors; no dedicated global interpersonal-security item was collected.

coefficients directly interpretable on the original 1–7 response scale. Third, composite scores are deterministic transformations of the raw item responses and can therefore be reproduced exactly from the original data by any researcher, without dependence on estimation software, identification constraints, or scoring methods. This transparency constitutes an additional methodological virtue in applied work where replicability is a priority. For these reasons, SaRI and SeRI were entered into the main analyses as manifest composite indices.

Then, for each transport option (i.e., manipulated route, not manipulated route, RSS), a third single global item asked, “*Would you feel safe/secure travelling on...?*” on a 1–7 scale, is used as an overall indicator and for convergent validity with the multi-item indices, consistent with whole-journey perspectives on transit safety/security. The three global “feel safe/secure” items were retained as standalone outcomes and as convergent-validity anchors expected to correlate inversely with the corresponding risk indices, in line with evidence that perceived safety/security shapes transit use (Newton, 2004). Table 2 details the items used to measure safety and security concepts and is self-explanatory.

As for the secondary outcomes, participants estimate the probability of inspection (i.e., Probability Estimates) for each route (manipulated and standard) (0–100%), thereby capturing hazard perception. Next, the preferred private mode of transportation (among walking, biking, scooter/gyropod, or car) and frequency of public transportation use (assessed using 7-points Likert scales from 1 meaning “almost never” to 7 meaning “almost every day”), fare evasion history (assessed on 7-points Likert scale ranging from 1 meaning “never” to 7 meaning “every time”), ticket purchase habits (whether they have single tickets or passes), waiting time (in minutes), and sociodemographic (age, gender, employment, education) are collected to serve as covariates and descriptive controls.

2.2. Pretesting and piloting

2.2.1. Pre-testing phase

Before launching the full experiment, a pretesting phase was conducted with a convenience (internal) sample to refine and validate the survey instrument. This stage aimed to identify overly complex or ambiguous questions and ensure the internal coherence of the experimental flow. Accordingly, we slightly revised the wording and coding of some items and defined all procedures for conducting a pilot test.

2.2.2. Pilot testing and power analysis

The pilot test, implemented after pretesting, served both to estimate the appropriate sample size and to assess the validity of the construct (or the reliability of the items) used to measure perceptions of safety and security. A group of students ($N_{users} = 37$, not included in this study) was recruited to participate in the pilot phase, enabling preliminary testing and refinement of the instrument. Based on pilot data, we computed the effect size ($ES = 0.8$, considered to be large according to standards, Cohen, 1988) by comparing transportation mode choices in the BSL condition.

Next, the following parameters ($ES = 0.8$, $\alpha = 0.05$, $power = 0.95$) are adopted to calculate the minimum sample size based on the Chi-Square Goodness-of-fit test (Faul et al., 2007). The required sample size turned out to be 33 independent observations per condition, for a total of 231. To account for it, we projected collecting 280 independent observations on aggregate (accounting for potential dropouts from participants who may fail the manipulation and attention-check questions).

2.3. The survey

2.3.1. Survey preparation

The overall experiment was programmed using PsyToolkit, an open-access web-based platform for behavioural research, to ensure a structured and interactive format (Stoet, 2010, 2016). Decisions are made within the scenario’s constraints; choices are single-shot and consequential only within the narrative (no repeated games, no possibility to go back or to modify one’s response). The protocol uses conditional branching logic, so the scenario and follow-up items adapt to the experimental group.

Ethical approval

Once the piloting was completed, and before administering the survey, the complete protocol, including materials, items, and parameterisation, was submitted to the host institution’s independent ethics committee for review and approval. Only after ethical clearance was granted, we proceeded to pre-registration. We displayed our research design and hypotheses publicly on osf.io prior to data collection in line with best standard practices in open sciences (Brodeur et al., 2024; Logg and Dorison, 2021) (<https://osf.io/rpf9w/overview>).

2.3.3. Survey administration

When considering how to administer the survey, several data-collection methods were evaluated, including paper-and-pencil interviews, computer-assisted telephone surveys, and online questionnaires. Each has its own advantages and disadvantages. For this study, an online survey was chosen for three main reasons. First, it enables the participation of many respondents while keeping operational costs to a minimum (e.g., interview time, data entry, and processing). Second, it is non-intrusive and requires limited effort from participants, making it easier to fit into their daily routines (Too and Earl, 2010). Third, by eliminating direct interviewer–respondent contact, self-administered online surveys attenuate social desirability pressures and elicit more candid reporting on socially sensitive behaviours such as fare evasion (Kaplan et al., 2012; Tourangeau and Yan, 2007). However, online surveys have certain limitations, such as underrepresenting some users who have limited Internet access. Despite this drawback, the questionnaire’s

length and complexity rendered an intercept survey impractical.

Respondents were recruited from the population of business school students in France, and recruitment invitations were distributed online to these students, with weekly reminders dispatched over a four-week period to maximise response rates. The questionnaire remained accessible for participation exclusively within this one-month timeframe. Data collection occurred in two waves, in April and September 2024. To ensure data integrity and prevent duplicate entries, participation was restricted to a single instance per individual; the software recorded IP addresses in a separate document to enforce this limitation.

2.3.4. Data preparation

At the end of the second data-collection wave, we closed the questionnaire and extracted the data from Psytoolkit. All responses are entered into the statistical software Stata (v.16.1), which was used to perform all data analyses.

We exclude data from participants failing both the attention and manipulation check questions. We prepare the data for analysis by coding treatment factors (Inspection vs. Fine, Day vs. Night), keeping BSL and Irrelevant conditions as comparison arms. We construct binary outcomes to choose the manipulated route over the alternatives. Finally, we construct SaRI and SeRI for each transportation mode (bus or RSS) as detailed in [section 2.1.2](#).

2.4. Data analysis method

A two-step approach is adopted to analyse the data. First, descriptive statistics are used to characterize the sample and assess the survey's robustness. Second, inferential statistics are applied to address the research questions and draw conclusions.

2.4.1. Descriptive analysis

Following the sample description, we measure the reliability of the Item by the Cronbach's alpha (α_c) ([Cronbach, 1951](#)). Next, we assess data pooling across waves and internal consistency to rule out experimental artefacts or salience effects. Analyses then examine transport choices under varying enforcement/contextual conditions, followed by fare-evasion behaviours and safety/security perceptions. This progression, from design robustness to behavioural/perceptual outcomes, clarifies the effects of enforcement on choices and safety across day/night.

In addition to internal consistency, we further stress the validity of the SaRI and SeRI indices through correlations with global items and a two-factor Confirmatory Factor Analysis (CFA). The CFA is used as a robustness check to verify whether the substantive conclusions would be sensitive to alternative measurement specifications.

2.4.2. Inferential analysis

This study tests four research questions (RQ1–RQ4) linking enforcement cues (inspection vs fines), perceived safety and security, travel choices (manipulated bus, standard bus, RSS), and fare-evasion behaviour (travelling without ticket, travelling without validating a ticket), with explicit day–night contrasts. The analytical strategy integrates deterrence theory and behavioural transport choice to examine both attitudinal responses (perceptions) and behavioural outcomes (route choice, fare evasion).

Two complementary estimation strategies are employed throughout the inferential analyses, each aligned with the measurement scale of the outcome variable.

Ordinary least squares (OLS) regression is applied to continuous or near-continuous perception outcomes, specifically, route-level safety differentials (Δ Safety, Δ Safety_ITT), the Safety Risk Index (SaRI), the Security Risk Index (SeRI), and perceived RSS safety (Safety_RSS). These variables are constructed as arithmetic means of 7-point Likert-type items; treating multi-item composite scores of this kind as approximately interval-level is a well-established and widely defended practice in the behavioural and social sciences ([Norman, 2010](#); [Sullivan and Artino, 2013](#)).

Table 3

Overview of analyses conducted according to each research question (RQ).

Analysis	RQ	Conditions	Sample	Method	Outcome(s)
#1	RQ1	Inspection vs. Fine	PP	OLS	Δ Safety
#2	RQ1	All enforcement arms ^a	PP	OLS	SaRI, SeRI
#3	RQ1	All arms ^b	ITT	OLS	Δ Safety_ITT, Safety_RSS
#4	RQ2	All arms ^b	ITT	Chi-square / proportions	Route choice
#5	RQ2	Inspection vs. Fine vs. Irrelevant	GSEM	GSEM	Route choice via perceptions
#6	RQ3	Inspection vs. Fine	PP	OLS with Arm $\times$ Period	Δ Safety, SaRI, SeRI
#7	RQ3	All arms ^b	GSEM	GSEM (Night as predictor)	Perception equations
#8	RQ4	Inspection vs. Fine	GSEM	GSEM	Light evasion, hard evasion
#9	RQ4	Inspection vs. Fine	GSEM	GSEM + Δ ProbInspect	Light evasion, hard evasion

Note. PP = Per-Protocol (restricted to participants assigned to enforcement arms: Inspection and Fine). ITT = Intention-To-Treat (all randomised participants). GSEM = Generalised Structural Equation Model. The GSEM sample ($N = 217$) comprises ITT participants ($N = 268$) minus those assigned to the Baseline condition ($n = 42$; no route-specific stimuli) and those who chose the ridesharing service ($n = 9$). Δ Safety = centred difference in perceived safety between the manipulated and non-manipulated route. SaRI = Safety Risk Index. SeRI = Security Risk Index. Δ ProbInspect = centred perceived difference in inspection probability between the two bus routes. "All enforcement arms" refers to all four per-protocol cells (INSP_D, INSP_N, FINE_D, FINE_N); Baseline and Irrelevant conditions are excluded by definition from the PP sample. "All arms" encompasses all seven experimental conditions (BSL, INSP_D, INSP_N, FINE_D, FINE_N, IRR_D, IRR_N).

Discrete behavioural outcomes (selection of the manipulated route, light fare evasion, and hard fare evasion) are binary indicators and are modelled using logistic estimators. Within the main analytical framework, these logit equations are estimated jointly with the linear perception equations via Generalised Structural Equation Models (GSEM), which allow simultaneous estimation of linear and logistic links while accounting for shared predictors and mediating pathways. As confirmatory robustness checks, standalone binary logit and multinomial logit specifications for the behavioural outcomes are estimated.

Building on this modelling framework, we next translate these estimation principles into an analytical structure capable of addressing the four research questions. Specifically, we implemented a set of pre-specified analyses tailored to the specific constructs and identification strategies relevant to each RQ. Therefore, each RQ is examined through one or more complementary analyses, explicitly identified as Analyses #1–#9, combining descriptive, causal, and multivariate modelling approaches.

To facilitate navigation of the empirical strategy, Table 3 provides a structured overview of the nine analyses, whose results are reported in Section 3. Each row maps a numbered analysis to its research question, experimental conditions, analytical sample, primary method, and outcome variable(s). Analyses #1–#3 address RQ1 (perceived safety and security); Analyses #4–#5 address RQ2 (route choice); Analyses #6–#7 address RQ3 (day/night moderation); and Analyses #8–#9 address RQ4 (fare evasion). The content and specification of each analysis are described in detail below, organised by research question.

First, RQ1 examines whether inspection or fine cues increase perceived safety and security through three distinct analyses. The analyses examine whether increasing inspection frequency or fine severity improves passengers' perceived safety and security, both at the level of specific bus routes and at the system level (Bus + RSS). For route-level outcomes, we construct contrasts such that positive differentials capture relatively safer evaluations of the manipulated route compared to the within-scenario control. The analyses are as follows. A per-protocol (PP) analysis (analysis #1) is conducted, namely an analysis restricted to participants assigned to the enforcement arms (Inspection and Fine), excluding Baseline and Irrelevant controls. Adherence is defined using prespecified criteria to estimate efficacy among compliers rather than to assess intent-to-treat effects. OLS regressions compare Fine versus Inspection on the centred safety differential, ΔSafety , defined as perceived safety on the manipulated route minus the non-manipulated route; positive values indicate relative safety gains for the manipulated route. Parallel models (namely A and B) use operational safety (SaRI) and security (SeRI) risk indices as robustness checks. Specifications include $\text{arm} \times \text{period}$ (Day vs. Night) interactions with heteroskedasticity-robust standard errors; Model A uses only design factors, while Model B adjusted for age, gender, public transport usage, and fare evasion history to enhance precision without biasing causal estimates.

An intention-to-treat (ITT) analysis (analysis #2) encompasses all randomised participants, and summarised treatment effects at the level of randomisation cells by comparing mean Safety_RSS scores across arms and periods using OLS with the same Model A / Model B structure as above. This test network-wide reassurance from deterrence cues to compare means, standard errors, and 95% confidence intervals by arm-period cell. Safety_RSS is prioritised because it uniformly measures system perceptions, complementing route-specific outcomes that are unavailable in Baseline.

We construct an ITT differential, $\Delta\text{Safety_ITT}$, defined as the perceived safety of the manipulated route minus that of its within-scenario control, irrespective of the option ultimately chosen. This differential is centred and modelled as a function of intervention arm, period and their interaction, using the same OLS framework (the Baseline arm was excluded because it did not feature route stimuli).

A further ITT analysis (analysis #3) targets route-specific safety in Inspection, Fine, and Irrelevant arms. We construct an ITT differential, $\Delta\text{Safety_ITT}$, defined as the perceived safety of the manipulated route minus that of its within-scenario control, irrespective of the option ultimately chosen. This differential is centred and modelled as a function of intervention arm, period and their interaction, using the same OLS framework (the Baseline arm was excluded because it did not feature route stimuli). We use the same model specification as above: Model A provides unadjusted design-based contrasts; Model B adds covariates for robustness. Equivalence tests using two one-sided tests (TOST) supplement all analyses to assess minimal effects (Lakens, 2017).

RQ2 examines whether anti-fare evasion strategies, by enhancing perceived safety and security, increase public bus use, particularly the selection of the manipulated route. It is examined through two complementary analyses. A descriptive analysis (analysis #4) summarises route choices across experimental arms (Baseline, Inspection, Fine, Irrelevant) and periods (Day vs. Night), reporting proportions selecting the manipulated bus, standard bus, or ridesharing service (RSS). Exact binomial tests compare manipulated versus standard route shares within arms, providing a design-based overview of choice patterns and manipulated route dominance. This step establishes empirical regularities prior to causal modelling. For analysis #5, we estimate a GSEM on participants who chose to return home by bus (manipulated or standard). This model links four centred perception measures ($\Delta\text{Safety_c}$, SaRI, SeRI, Safety_RSS), modelled linearly, to three key behaviours modelled logistically (participant chooses the manipulated route, engages in light fare evasion, engages in hard fare evasion). Predictors include enforcement arm (with Inspection as reference), travel period (night vs. day), and centred individual covariates (age, gender, public transport usage, fare-evasion history). This multivariate specification is well suited to RQ2 because it (i) exploits the randomised arms for identification of upstream effects, (ii) allows simultaneous estimation of direct and indirect (mediated) paths, and (iii) avoids fitting separate models for each outcome while accounting for shared covariates and mediators. In the GSEM, SaRI and SeRI are treated as observed composite indices rather than latent variables. These indices, ΔSafety , and Safety_RSS are mean-centred before entering the model. The GSEM combines linear equations for the four perception outcomes with logit equations for the three discrete behavioural outcomes. Specifically, the enforcement arm, travel period and centred individual covariates predict the perception measures, which in turn enter the behavioural equations predicting the choice of the manipulated route, light fare evasion and hard fare evasion. This structure reflects the dual pathway of deterrence theory: the Inspection treatment operationalises a cue related to the probability of detection, whereas the Fine treatment operationalises sanction severity (Becker, 1968; Polinsky and Shavell, 2000). Therefore, the model tests whether enforcement cues affect behavioural outcomes both directly and indirectly through perceived safety, perceived security and route choice. In the baseline specification, the perception

equations are estimated jointly with a common set of predictors, but residual covariances among perception equations are not freely estimated. This simplifying assumption is complemented by the correlation diagnostics. Moreover, the key behavioural models are re-estimated with CFA-derived factor scores substituted for the composite indices, allowing us to examine the stability of the results under an alternative measurement approach.

RQ 3 asks whether night-time travel, compared with day-time, shifts safety and risk perceptions across anti-fare-evasion arms and alters the strength with which these perceptions influence the choice of the manipulated route among treated participants. It is analysed within two complementary analyses. Analysis #6 consists of an OLS regression, restricted to the treated sample (Inspection and Fine arms), centred and regressed safety/risk perceptions (SafetyLA, SaRI, SeRI) on the Arm $\times$ Day/Night factor. Centring and explicit interaction terms follow standard recommendations for probing moderation in regression and provide a descriptive check on whether the experimental factors and the night condition shift perceived safety levels before modelling their impact on choice. We use the same model specifications as for RQ1 (Model A / Model B). Predicted margins assessed overlapping intervals across cells, checking if night differentially alters perceptions prior to choice modelling.

Analysis #7 departs from the GSEM developed for RQ2 to examine period main effects and period $\times$ perception interactions on route choice among treated participants. The full model incorporates Night (vs. Day) throughout, enabling tests of nocturnal amplification in perception-choice links via interaction paths and stratified estimates. This multivariate approach converges with OLS, preserving causal identification while accounting for mediators and correlates.

Finally, RQ4 examines whether enforcement information (inspection versus fine versus irrelevant information) and perceived inspection probabilities shape fare evasion behaviour over and above the characteristics of the bus routes themselves. Therefore, we implement two analyses. First (analysis #8), we leverage the primary generalised structural equation model (GSEM; Analysis #5), from RQ2 by focusing on its light (travelling with a non-validated ticket) and hard (travelling without a ticket) fare evasion equations. The GSEM regresses binary evasion outcomes on enforcement arms (Inspection reference; Fine, Irrelevant), period (Night vs. Day), perceptions (ΔSafety^c , SaRI, SeRI, Safety_RSS), route choice, and centred covariates (age, gender, transport usage, evasion history). Direct arm effects, plus indirect paths via perceptions/choice, test enforcement impacts net of line characteristics and confounders.

A complementary robustness GSEM (analysis #9) extends this joint system across all arms and outcomes, and introduces the centred perceived difference in inspection probabilities between the two bus routes ($\Delta\text{ProbInspect}$). Enforcement and period predict perceptions (including $\Delta\text{ProbInspect}$), which mediated effects on choice and evasion (light/hard). This specification isolates belief shifts from Fine cues while assessing downstream behavioural links, using logit links and robust standard errors.

As a general principle across all four RQs, and to aid interpretation beyond formal statistical significance while safeguarding against the misinterpretation of null findings, we evaluate effect-size indices adapted to each model family applied across the entire analytical framework. Specifically, we compute partial Cohen's f^2 for focal terms in linear models (OLS), Cohen's w for chi-square associations involving categorical choices, and Odds Ratios (OR) for logistic equations within the GSEM frameworks. Where useful, conventional logit-to- d transformations are assessed as supplementary interpretative benchmarks to evaluate the practical magnitude of the estimated parameters.

3. Results

This section describes the sample and robustness checks, then presents findings on transportation choices, and finally reports results addressing each research question.

3.1. Sample description

The ITT analytical sample comprised 268 participants, who were on average 19.68 years old ($SD = 1.68$), predominantly female (57.84%), and almost exclusively students (98.13%). Participants reported frequent public transport use ($M = 5.50$, $SD = 1.34$). The PP analytical sample comprised 146 participants and showed a similar profile, with a mean age of 19.56 years ($SD = 1.61$), 65.75% female participants, 98.63% students, and high public transport use ($M = 5.53$, $SD = 1.29$). The sample is described in [Supplementary Material – Appendix C](#).

3.2. Measurement validity of perception indices

Internal consistency ranged from acceptable (bus SaRI, $\alpha = 0.6484$) to robust (bus SeRI, $\alpha = 0.9003$; RSS indices, $\alpha = 0.8360$ and 0.8849 ; [Table 4](#)).

Correlations among SaRI, SeRI and the route-level global safety ratings were small and statistically non-significant (all $|r| \leq 0.097$,

Table 4
Reliability statistics to measure safety and security perceptions.

Scale	Items	N	Cronbach's α_c	Notes on α_c
Bus – SaRI	0–5	268	0.6484	Acceptable
Bus – SeRI	6–10	268	0.9003	Excellent
RSS – SaRI	0–5	268	0.8360	Robust/High
RSS – SeRI	6–10	268	0.8849	Robust/High

all $p \geq 0.11$; [Supplementary Material – Appendix D](#) –Table D1). These near-zero associations are expected: SaRI and SeRI capture perceived probabilities of specific enumerated hazards, whereas the global safety ratings reflect broader holistic evaluations of a given route. Therefore, these two measures operate at different levels of abstraction within the safety-perception domain, and the near-zero correlations are interpreted as a discriminant validity check rather than a convergent validity failure.

The weaker consistency of the bus-mode SaRI, also reflected in the more heterogeneous standardised loadings of the operational safety-risk factor in the CFA (range: 0.26–0.77; [Supplementary Material – Appendix D](#) –Table D2), is acknowledged as a limitation of that composite. The bus-mode indices were moderately positively correlated ($r = 0.484$, $p < 0.001$), indicating that respondents who perceived greater operational safety risks also tended to perceive greater interpersonal security risks, while each index retained approximately 77% unique variance. CFA-based robustness analyses, in which composite indices were replaced by factor scores, yielded substantively unchanged behavioural conclusions. CFA fit statistics and factor-score robustness checks are reported in [Supplementary Material – Appendix D](#) –Tables D3 and D4.

The construct-validity and measurement-robustness checks reported in this section are restricted to SaRI and SeRI because these are the only variables in the study operationalised as multi-item psychometric scales. Internal-consistency reliability diagnostics (Cronbach's α) and CFA are methodologically meaningful only when several items are designed to measure a common underlying construct ([Cronbach, 1951](#); [Taber, 2018](#)); single-item measures cannot, by definition, yield internal-consistency estimates. The remaining perception variables, namely the route-level global safety ratings (Safety_LA, Safety_LB, Safety_RSS) and the derived route-safety differential (Δ Safety), are single-item or arithmetic difference scores and therefore do not require latent-variable robustness checks. Similarly, the behavioural outcomes (route choice, light and hard fare evasion) are observed binary indicators rather than psychometric latent constructs ([Diamantopoulos et al., 2012](#)). For these variables, robustness was assessed through the modelling strategy itself, specifically through the standalone binary logit and multinomial logit specifications reported in [Supplementary Material – Appendix E](#) –Tables E1, E2, and E3, rather than through psychometric reliability diagnostics.

3.3. RQ1 – do fare-evasion deterrence strategies improve perceived safety and security?

RQ1 was examined through three complementary analyses.

Analysis #1 focused on the PP sample and compared the two enforcement arms, Inspection and Fine, among participants who met the predefined inclusion criteria. This PP analysis included 146 participants and used OLS models to estimate the effect of enforcement type on the centred route-specific safety differential, Δ Safety, defined as perceived safety on the manipulated route minus perceived safety on the standard route. Positive values therefore indicate relative safety gains for the manipulated route. Across both the un-adjusted specification, Model A, and the covariate-adjusted specification, Model B, Fine–Inspection contrasts were small and statistically uncertain, with confidence intervals spanning zero. The Arm $\times$ Period interaction term, comparing Night and Day conditions, likewise remained close to null. Parallel PP models using the operational SaRI, and the SeRI, as alternative outcomes showed the same substantive pattern: the Fine and Inspection treatments did not differ meaningfully in their effects on perceived safety or security ([Tables 5 and 6](#)).

Effect-size estimates reinforce this interpretation. Across the PP specifications, partial Cohen's f^2 values for the focal enforcement contrast ranged from 0.003 to 0.018, all below the conventional threshold for a small effect, $f^2 = 0.02$ ([Cohen, 1988](#)). Substantively, these values indicate that the distinction between Fine and Inspection accounted for a negligible share of between-participant variation in perceived safety and security once the model predictors were considered. The largest observed standardised contrast across specifications was approximately Cohen's $d = 0.27$, which falls in the small-to-modest range. This estimate should be interpreted as the largest observed point estimate rather than as an upper confidence-limit bound. Thus, although the confidence intervals do not allow

Table 5
OLS models of safety and risk perceptions under Fine versus Inspection (per-protocol sample).

Dep. Var	Δ Safety (centred)		SaRI		SeRI	
Model	Model A	Model B	Model A	Model B	Model A	Model B
Fine	0.322(0.219)	0.310(0.222)	0.290 (0.187)	0.295 (0.186)	0.287 (0.333)	0.194 (0.295)
Night	0.120(0.286)	0.077(0.275)	0.107 (0.173)	0.135 (0.175)	0.116 (0.312)	0.182 (0.280)
Interaction (Fine $\times$ Night)	−0.108(0.326)	−0.058(0.327)	−0.117 (0.261)	−0.132 (0.264)	−0.137 (0.458)	−0.097 (0.414)
Age	–	0.039(0.053)	–	−0.031 (0.041)	–	−0.020 (0.061)
Female	–	−0.322(0.205)	–	0.382*** (0.129)	–	1.291*** (0.195)
Public Transport Usage	–	0.243** (0.121)	–	0.018 (0.047)	–	0.012 (0.087)
Fare Evasion History	–	−0.124* (0.052)	–	0.055 (0.044)	–	0.036 (0.072)
Constant	−0.184(0.191)	0.043(0.194)	−0.160(0.114)	−0.424*** (0.140)	−0.158(0.214)	−1.008*** (0.216)
R ²	0.020	0.090	0.025	0.088	0.008	0.207
N	146	146	146	146	146	146

Note: Ordinary Least Squares estimates for centred global safety differences between manipulated and standard routes (Δ Safety_c) and for centred bus-level operational (SaRI_c) and security (SeRI_c) risk indices, among treated participants (Inspection and Fine arms only). Model A includes Fine (vs Inspection), Night (vs Day), and the Fine $\times$ Night interaction. Model B adds centred age, gender (Female), public transport usage frequency, and fare evasion history as covariates. Robust standard errors are reported in parentheses. All continuous covariates are mean-centred on the per-protocol sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6Adjusted means for Δ Safety / SaRI / SeRI by arm and period.

Outcome	Period	Arm	Estimate	95% CI	p-value
Δ Safety	Day	Inspection	-0.189	[-0.527, 0.150]	0.272
	Day	Fine	0.133	[-0.163, 0.428]	0.378
	Night	Inspection	-0.081	[-0.414, 0.253]	0.633
	Night	Fine	0.155	[-0.042, 0.352]	0.121
SaRI	Day	Inspection	-0.162	[-0.389, 0.064]	0.159
	Day	Fine	0.133	[-0.153, 0.418]	0.360
	Night	Inspection	-0.048	[-0.299, 0.203]	0.704
	Night	Fine	0.114	[-0.138, 0.366]	0.372
SeRI	Day	Inspection	-0.118	[-0.485, 0.248]	0.525
	Day	Fine	0.076	[-0.368, 0.521]	0.735
	Night	Inspection	-0.016	[-0.435, 0.402]	0.939
	Night	Fine	0.081	[-0.335, 0.496]	0.702

us to rule out all effects of practical relevance, the observed pattern does not suggest a substantively large reassurance effect of either enforcement strategy.

These findings should nevertheless be interpreted considering statistical power. With $N = 146$ in the PP sample, the models had limited sensitivity to detect effects below the small-effect benchmark. Under a one-degree-of-freedom focal-contrast approximation, achieved power was approximately 0.40 for detecting $f^2 = 0.02$ at $\alpha = 0.05$. In other words, the study had roughly a 40% chance of detecting a genuine small effect, had one existed. Achieving 80% power to detect an effect as small as the largest observed partial f^2 value, $f^2 = 0.018$, would require a substantially larger sample of approximately $N > 400$, depending on the exact model specification (Faul et al., 2007). Therefore, the null findings should not be read as definitive evidence that enforcement cues have no effect. Rather, they indicate that Fine–Inspection differences were empirically small in this study and estimated with limited precision.

Analysis #2 extended the assessment to the intention-to-treat sample and examined network-level reassurance effects using perceived RSS safety (Safety_RSS) across randomised participants. The ITT OLS models provided no consistent evidence that enforcement conditions improved perceived safety relative to the comparison conditions. Descriptive means by condition and time of travel are reported in Table E4, while the corresponding ITT OLS models are reported in Table E5 (see [Supplementary Material – Appendix E](#)).

Analysis #3 examined the ITT route-specific safety differential (Δ Safety_ITT) among the arms for which route-level stimuli were available. This analysis converged with the PP results: enforcement cues did not produce a robust route-level improvement in perceived safety, and the predicted margins for Δ Safety_ITT remained close to zero with confidence intervals spanning zero across conditions ([Supplementary Material – Appendix E](#) – Tables E5 and E6).

Supplementary equivalence tests using Two One-Sided Tests, TOST, are also reported in Table E7 ([Supplementary Material – Appendix E](#)). They did not establish equivalence within small-effect bounds, meaning that the analyses cannot formally demonstrate that the enforcement effects are practically zero. Rather, the TOST results highlight the same limitation as the power analysis: the evidence points to small and statistically uncertain effects, but the study was not sufficiently powered to make strong equivalence claims.

Taken together, Analyses #1 to #3 provide little evidence that increasing inspection frequency or raising fines meaningfully improves passengers' perceived safety or security in this vignette experiment. The null enforcement pattern is accompanied by a more systematic role of gender across specifications: women reported higher perceived safety and security risks than men throughout, suggesting that individual sociodemographic characteristics may be stronger drivers of safety and security perceptions than marginal variations in enforcement intensity.

3.4. RQ2 – do enforcement cues affect route choice through safety and security perceptions?

Analysis #4 examined the descriptive distribution of choices across experimental conditions. Pooling across conditions ($N = 268$), the manipulated route was the dominant choice, selected by 63.43% of participants, compared with 32.84% for the standard route and 3.73% for the RSS alternative (Table 7). Exact tests of proportions indicated that the manipulated route was chosen significantly more often than the standard route in all conditions except Fine, where the two shares were statistically indistinguishable ($p = 0.492$). The very low RSS uptake indicates that most observed choice variation occurred between the two bus alternatives rather than between bus

Table 7

Bus and RSS choices across conditions.

Condition	RSS	Standard Bus route	Manipulated Bus route
BSL	4.76%	28.57%	66.67%
Inspection Conditions (INSP_D + INSP_N)	1.28%	34.62%	64.10%
Fine Conditions (FINE_D + FINE_N)	5.88%	44.12%	50.0%
Irrelevant Conditions (IRR_D + IRR_N)	3.75%	23.75%	72.5%
Aggregated Data	3.73%	32.84%	63.43%

and ride-sharing.

Analysis #5 examined whether route choice was shaped by safety and security perceptions, and whether enforcement cues affected choice directly or indirectly through these perceptions. The main GSEM was estimated on the complete-case subsample of participants who selected one of the two bus routes and had non-missing values on the variables entering the model ($n = 217$). The model jointly estimated the perception equations for Δ Safety, SaRI, SeRI and Safety_RSS, together with a logit equation predicting selection of the manipulated route. Noteworthy, in this empirical setting, specifying latent measurement models for SaRI and SeRI would have added roughly 23 free parameters (nine loadings, eleven error variances, two factor variances, one covariance), which is disproportionate relative to the available sample and would reduce estimation precision in the behavioural equations. Composite indices were therefore retained in the main GSEM. Structural coefficients are reported in Table 8, and the model is summarised in Fig. 2.

Within the perception block, enforcement arm and travel period showed only modest and statistically uncertain associations with the four perception outcomes, consistent with the RQ1 finding that deterrence cues did not reliably reshape safety or security perceptions. Female participants reported significantly higher operational safety risk (SaRI) and security/disorder risk (SeRI), and self-reported fare evaders perceived the RSS as comparatively safer.

In the route-choice equation, the perceived safety differential between the manipulated and standard routes was a robust positive predictor of choosing the manipulated route, $OR = 1.866$, 95% CI [1.259, 2.740], $p = 0.002$. Therefore, participants who perceived the manipulated route as comparatively safer were more likely to select it. By contrast, SaRI, SeRI and Safety_RSS were not significant predictors once Δ Safety was included. Female participants were also marginally more likely to select the manipulated route. Overall, these results indicate that route choice was more strongly predicted by participants' comparative safety perception of the manipulated route relative to the standard route than by broader modal- or system-level appraisals of safety and security.

Enforcement cues provided no meaningful empirical support for an indirect perceptual pathway to route choice. The Fine and Irrelevant arms showed only weak and imprecise associations with the perception outcomes, and formal indirect effects through the perception measures were essentially zero. For example, the indirect effect of Fine versus Inspection through Δ Safety was approximately 0.000 on the log-odds scale, with a confidence interval close to $[-0.001, 0.001]$; the indirect effect through Safety_RSS was similarly negligible, approximately 0.001 with a confidence interval close to $[-0.001, 0.003]$. The hypothesised mediation pathway from enforcement cues to route choice through safety and security perceptions is therefore not supported.

The GSEM nevertheless indicated a direct negative association between the Fine arm and selection of the manipulated route relative to Inspection, $OR = 0.456$, 95% CI [0.220, 0.955], $p = 0.037$. This estimate suggests that fine-based messaging may have reduced the attractiveness of the manipulated route through a channel not captured by the measured safety and security perceptions, possibly because sanction-focused messages can be interpreted as coercive or punitive rather than reassuring (Fürst and Herold, 2018; Suquet, 2010). However, the confidence interval was close to unity and reduced-form robustness checks yielded only marginal evidence for the same contrast. Therefore, this direct Fine effect should be interpreted cautiously.

Supplementary robustness checks using binary logit and multinomial logit specifications are reported in Supplementary Material – Appendix E – Tables E1 and E2. These models reproduced the same directional pattern, namely a lower uptake of the manipulated route under Fine than under Inspection, but the contrast was only marginally significant. Global model fit was small across all reduced-form specifications, with McFadden pseudo- R^2 ranging from 0.027 to 0.054 and corresponding Cohen's f^2 values, computed as $R^2/(1 - R^2)$, ranging from 0.028 to 0.057. These values confirm that the enforcement arm explained only a modest share of overall choice variability.

Taken together, Analyses #4 and #5 provide partial support for RQ2. Perceived comparative safety strongly predicted route choice, but enforcement cues did not meaningfully shift route choice through the measured safety and security perceptions. The only direct enforcement-related pattern was a negative Fine-versus-Inspection association with choosing the manipulated route, but this effect was fragile across specifications and should therefore be interpreted as suggestive rather than definitive.

3.5. RQ3 – does night-time travel strengthen the role of safety and security perceptions?

Analysis #6 tested whether the travel period moderated the effect of enforcement cues on safety and security perceptions among participants assigned to the Inspection and Fine arms. OLS models incorporating Arm $\times$ Period interaction terms were estimated for three perception outcomes: the safety differential (Δ Safety), the Safety Risk Index (SaRI), and the Security Risk Index (SeRI). Across both the unadjusted specification (Model A) and the covariate-adjusted specification (Model B), the interaction terms were small and statistically uncertain, with partial Cohen's f^2 values ranging from 0.0002 to 0.0018, uniformly below the conventional threshold for a small effect ($f^2 = 0.02$; Cohen, 1988). The main effect of night-time travel on perceived safety and security was similarly small and imprecisely estimated. Predicted margins showed highly overlapping confidence intervals between daytime and night-time conditions across both enforcement arms and all three perception outcomes. Detailed coefficients are reported in Table 5, and the corresponding predicted margins are reported in Table 6.

Analysis #7 provided a convergent structural assessment within the GSEM estimated for RQ2 (Analysis #5; Table 8). In this model, Night was included as a direct predictor in the perception and route-choice equations rather than as a formal interaction term. Consistent with the OLS findings, the Night coefficient was at most modest and imprecisely estimated across the perception equations, including Δ Safety, SaRI, SeRI, and Safety_RSS. The route-choice equation likewise provided no indication that night-time travel materially changed the selection of the manipulated route once the perception measures and experimental conditions were accounted for. Therefore, the structural model converged with the OLS finding that night-time conditions had no substantive direct effect on the perception outcomes, providing additional structural support for the absence of a nocturnal perceptual shift.

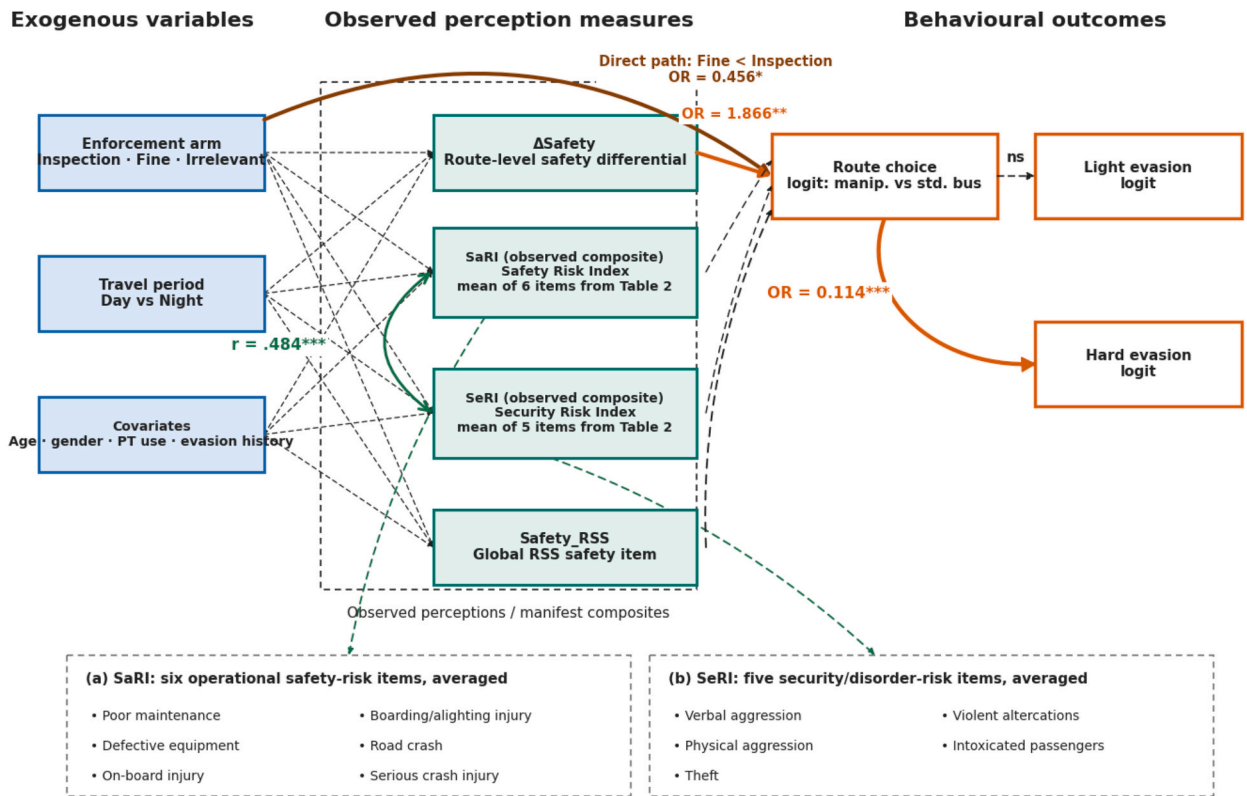
Taken together, Analyses #6 and #7 provide little support for RQ3. Night-time travel did not meaningfully amplify the effect of

Table 8

GSEM Full structural paths with robust standard errors.

Mediator (outcome)	Predictor (reference = Inspection)	Estimates	95% CI	p
Panel A – Perceptual mediators (Gaussian models: unstandardised coefficients β)				
Δ Safety	Fine vs Inspection	0.284 (1.328)	[-0.035, 0.603]	0.081
	Irrelevant vs Inspection	0.237 (1.267)	[-0.046, 0.519]	0.101
	Night vs Day	0.023 (1.023)	[-0.225, 0.270]	0.858
	Age	0.019 (1.019)	[-0.050, 0.087]	0.595
	Female	-0.203 (0.816)	[-0.490, 0.084]	0.166
SaRI	PT usage	0.136 (1.141)	[-0.034, 0.305]	0.116
	Past fare evasion	-0.062 (0.939)	[-0.144, 0.020]	0.140
	Fine vs Inspection	0.229 (1.257)	[-0.020, 0.478]	0.071
	Irrelevant vs Inspection	0.104 (1.109)	[-0.137, 0.346]	0.398
	Night vs Day	-0.006 (0.994)	[-0.213, 0.200]	0.952
SeRI	Age	-0.039 (0.961)	[-0.099, 0.021]	0.206
	Female	0.503 (1.653)	[0.300, 0.707]	<0.001
	PT usage	0.019 (1.019)	[-0.053, 0.092]	0.603
	Past fare evasion	0.058 (1.059)	[-0.019, 0.135]	0.137
	Fine vs Inspection	0.153 (1.165)	[-0.264, 0.571]	0.471
Safety RSS	Irrelevant vs Inspection	0.023 (1.023)	[-0.327, 0.374]	0.895
	Night vs Day	0.075 (1.077)	[-0.240, 0.389]	0.641
	Age	-0.030 (0.970)	[-0.133, 0.073]	0.565
	Female	1.328 (3.773)	[1.038, 1.618]	<0.001
	PT usage	-0.030 (0.970)	[-0.151, 0.091]	0.628
	Past fare evasion	0.034 (1.034)	[-0.086, 0.154]	0.582
	Fine vs Inspection	0.242 (1.273)	[-0.376, 0.860]	0.443
	Irrelevant vs Inspection	0.037 (1.037)	[-0.585, 0.659]	0.907
	Night vs Day	0.486 (1.625)	[-0.049, 1.022]	0.075
	Age	0.114 (1.120)	[-0.066, 0.294]	0.214
	Female	-0.157 (0.854)	[-0.713, 0.399]	0.581
	PT usage	-0.021 (0.979)	[-0.244, 0.201]	0.850
	Past fare evasion	0.337 (1.400)	[0.164, 0.510]	<0.001
Panel B – Route choice (logit models: odds ratios)				
Choosing the manipulated line (Choose_manip)	Fine vs Inspection	-0.785 (0.456)	[0.217, 0.955]	0.037
	Irrelevant vs Inspection	0.287 (1.333)	[0.635, 2.798]	0.448
	Night vs Day	0.503 (1.655)	[0.886, 3.092]	0.114
	Δ Safety	0.623 (1.866)	[1.259, 2.764]	0.002
	SaRI (bus)	0.042 (1.043)	[0.690, 1.578]	0.841
	SeRI (bus)	-0.029 (0.971)	[0.749, 1.260]	0.825
	Safety_RSS	-0.054 (0.947)	[0.813, 1.103]	0.481
	Age	0.041 (1.042)	[0.869, 1.250]	0.657
	Female	0.726 (2.067)	[0.979, 4.364]	0.057
	PT usage	0.035 (1.036)	[0.824, 1.301]	0.764
	Past fare evasion	-0.033 (0.967)	[0.759, 1.230]	0.782
Panel C – Light fare evasion (logit models: odds ratios)				
Light evasion (travelling with a non-validated ticket)	Fine vs Inspection	0.815 (2.261)	[0.598, 8.543]	0.229
	Irrelevant vs Inspection	0.929 (2.532)	[0.646, 9.922]	0.183
	Night vs Day	-0.929 (0.395)	[0.122, 1.287]	0.123
	Δ Safety	0.053 (1.055)	[0.348, 3.196]	0.925
	SaRI (bus)	-0.263 (0.768)	[0.419, 1.406]	0.392
	SeRI (bus)	0.474 (1.608)	[0.996, 2.596]	0.052
	Safety_RSS	0.052 (1.054)	[0.815, 1.362]	0.690
	Choose manipulated route	-0.208 (0.812)	[0.277, 2.384]	0.705
	Age	0.036 (1.037)	[0.777, 1.383]	0.805
	Female	-0.131 (0.877)	[0.208, 3.689]	0.857
	PT usage	-0.163 (0.849)	[0.563, 1.280]	0.434
	Past fare evasion	0.568 (1.765)	[1.323, 2.354]	<0.001
Panel D – Hard fare evasion (logit models: odds ratios)				
Hard evasion (travelling without ticket)	Fine vs Inspection	0.298 (1.348)	[0.261, 6.961]	0.721
	Irrelevant vs Inspection	0.400 (1.493)	[0.306, 7.291]	0.620
	Night vs Day	1.595 (4.933)	[1.386, 17.563]	0.014
	Δ Safety	-0.400 (0.670)	[0.331, 1.357]	0.266
	SaRI (bus)	-0.441 (0.643)	[0.292, 1.417]	0.274
	SeRI (bus)	-0.054 (0.947)	[0.555, 1.616]	0.841
	Safety_RSS	-0.211 (0.809)	[0.617, 1.062]	0.127
	Choose manipulated route	-2.171 (0.114)	[0.048, 0.269]	<0.001
	Age	-0.008 (0.992)	[0.686, 1.435]	0.967
	Female	-0.629 (0.533)	[0.154, 1.844]	0.320
	PT usage	0.212 (1.237)	[0.851, 1.797]	0.264
	Past fare evasion	0.524 (1.690)	[1.143, 2.498]	0.009

Notes. Panel A: coefficients; Panel B to D: Odds Ratios. Message condition = Inspection (visual enforcement), Fine (sanction-oriented enforcement), Irrelevant (neutral information control). “Enforcement” refers to contrasts between enforcement-oriented arms (Inspection $\cup$ Fine) vs Irrelevant, or to Fine vs Inspection. Gaussian panels report unstandardised coefficients (β) with 95% CIs; logit panels report odds ratios (OR) with 95% CIs (OR = exp (β)). Robust standard errors; N = 217 (bus sample). Continuous predictors centred. Intercepts omitted for brevity.



Note. Rectangles denote manifest/observed variables. SaRI and SeRI are composite indices formed by simple averaging of Table 2 items; they are not specified as latent factors in the main GSEM. The empirical SaRI–SeRI association is shown for transparency. Residual covariances among perception equations are constrained to zero in the baseline GSEM; correlation diagnostics and CFA-based robustness checks are reported in Appendix B. Direct paths from enforcement arm to evasion outcomes are estimated in the GSEM but omitted here for clarity; all were non-significant. Perception paths to evasion outcomes are similarly estimated but non-significant and omitted. Statistics displayed on paths are odds ratios (ORs); * $p < .05$, ** $p < .01$, *** $p < .001$.

teal = Observed measure - - - black = Estimated path orange = Significant perception path dark brown = Significant direct path green = Empirical correlation

Fig. 2. GSEM path diagram linking enforcement, observed perception measures, route choice, and fare-evasion outcomes.

enforcement cues on safety or security perceptions. The available evidence does not support the hypothesis that the travel period operated through the perceptual pathway tested in RQ3. However, as reported under RQ4, night-time travel was independently associated with fare-evasion behaviour in the GSEM evasion equations. Therefore, period effects appear behaviourally consequential, but through a mechanism distinct from the safety and security perception pathway examined here.

3.6. RQ4 – do enforcement cues reduce fare-evasion behaviour?

Analysis #8 examined whether fare-evasion behaviour, operationalised as stated evasion intentions in the vignette task, was shaped by the enforcement treatment. The analysis focused on the light and hard evasion equations of the primary GSEM reported in Table 8. In neither outcome did the Fine versus Inspection contrast reach statistical significance, and confidence intervals spanned unity. For light evasion, defined as travelling without ticket validation, the Fine arm was associated with an odds ratio above one in the primary GSEM, OR = 2.26, 95% CI [0.60, 8.54], indicating a direction inconsistent with a simple deterrence prediction. However, this estimate was imprecise and statistically non-significant. Reduced-form logit checks (Supplementary Material – Appendix E – Table E3) and the alternative GSEM (including the ΔProbInspect-augmented GSEM robustness specification estimated in Analysis #9, Table E8 (Supplementary Material – Appendix E) specifications produced the same substantive conclusion: Fine-versus-Inspection contrasts varied in magnitude, but none provided evidence of a compliance-inducing effect. For hard evasion, defined as travelling without a ticket, the Fine-versus-Inspection estimate was also non-significant in the primary GSEM, OR = 1.35, 95% CI [0.26, 6.96], in a context

where this outcome had a low base rate in the analytical sample. Taken together, these results provide no empirical support for a differential compliance-inducing effect of Fine relative to Inspection.

Against this null enforcement pattern, two predictors showed larger and more precise associations with evasion. Across the primary GSEM and alternative specifications, Night odds ratios for hard evasion ranged from 3.12 to 4.93, corresponding to standardised effect sizes of approximately $d = 0.63$ – 0.88 under the standard logit-to- d conversion (Chen et al., 2010). These estimates place the association in the medium-to-large range; even the lower-bound estimate remains substantively meaningful, although all estimates should be interpreted cautiously given the low number of hard-evasion events in the analytical sample, which limits precision across specifications. Individual fare-evasion history was a consistently strong and statistically significant predictor of light evasion across specifications and reached statistical significance in the primary GSEM hard-evasion equation.

This comparative pattern across effect magnitudes provides critical evidence for the experimental validity of the design. While the Fine versus Inspection contrast produced imprecise and non-significant ORs for evasion (OR = 2.26 for light evasion, OR = 1.35 for hard evasion in the primary GSEM), the GSEM equations robustly captured the night-time effect on hard evasion. Under standard logit-to- d conversion, this night effect translates to a substantial standardised mean difference (d between 0.63 and 0.88). The fact that the statistical system successfully isolates large, robust associations where they genuinely exist (such as night-time travel and individual evasion history) argues strongly against a purely 'weak-design' explanation, confirming instead that the nominal enforcement cues tested here produced small observed effects under ecologically plausible vignette conditions.

For robustness, Analysis #9 augmented the GSEM with the centred perceived difference in inspection probabilities between routes, $\Delta\text{ProbInspect}$; complete results are reported in [Supplementary Material – Appendix E](#) –Tables E8. The Fine condition robustly shifted $\Delta\text{ProbInspect}$, $\beta = -13.57$, $\text{SE} = 2.33$, $p < 0.001$, indicating that participants perceived a lower inspection-probability differential in the Fine condition. However, $\Delta\text{ProbInspect}$ did not significantly predict any downstream behavioural outcome in the GSEM (route choice, light evasion, or hard evasion), suggesting that the belief shift induced by the Fine condition had no detectable compliance effect.

4. Discussion & conclusion

This paper sets out to examine whether anti-fare evasion strategies (such as raising the number of ticket inspection operations and the level of fines) implemented by TAs/PTCs may influence passengers' travel choices on Public Transport (PT), compliance behaviour, and perceptions of safety and security, particularly in comparison to Ridesharing Services (RSS) and across day and night contexts. Using a vignette-based randomised experiment, we manipulated enforcement cues (increased inspections versus higher fines) and time of travel (daytime vs. night-time) across two bus routes to examine their combined effects on behavioural intentions and psychological perceptions. Four research questions examined whether anti-fare evasion strategies affect perceived safety, travel choices, and compliance.

Three main empirical findings emerge.

First, regarding perceived safety and security, enforcement severity did not meaningfully enhance these perceptions when contextual features were considered. Perceived safety and security appeared to be shaped primarily by route design, time of day, and modal context, rather than by the incremental severity of the enforcement message. This is consistent with the view that reassurance arises from visible and certain guardianship rather than higher nominal severity. This aligns with the empirical evidence showing that neither inspection nor fine cues produced statistically reliable improvements in SaRI, SeRI, or route-level safety differentials.

Second, passengers' route preferences were driven primarily by perceived safety and security and by modal context, rather than by enforcement type. Differences between inspection- and fine-based regimes were small and statistically uncertain, and night-time travel did not substantially modify these patterns. RSS exhibited a modest night-time safety advantage, in line with gendered and temporal patterns reported in the transit safety literature, and night travel did not materially change the safety and risk perceptions underlying route choice. These patterns are consistent with the finding that enforcement cues did not significantly shift route choice probabilities once perceptions were accounted for.

Third, fare evasion patterns were more strongly associated with time, route, and individual history of fare evasion than with enforcement intensity. Ticket purchase propensity was lower at night and varied across routes, whereas per-protocol models and two-sample proportion tests provided no statistically significant evidence that fines reduced either hard or light evasion relative to inspections once period, route, and covariates were controlled. This pattern is more consistent with theoretical accounts emphasising capable guardianship and visible control (rather than isolated increases in sanction severity) as the main levers for both perceived safety and compliance. The absence of detectable differences between Inspection and Fine conditions reinforces the limited behavioural impact of sanction severity in this setting.

All these findings suggest that compliance depends not only on detection probability or fine severity, but also on how travellers perceive safety and security, fairness, and social order in public transport environments. Overall, these findings challenge the assumption that deterrence measures improve perceived safety and highlight the need for integrated strategies that combine enforcement with broader initiatives to enhance security, particularly during night travel, to sustain public transport attractiveness in an increasingly multimodal environment.

4.1. Strengths of the study

This study advances the fare-evasion and safety and security literature by jointly analysing deterrence levers (increased inspections vs higher fines), time of travel (day vs night), and modal competition with a RSS alternative. That design connects the economics of

public enforcement, where policy levers operate through the *probability* and *severity* of fines, with behavioural responses and perceived guardianship (Becker, 1968; Polinsky and Shavell, 2000). The vignette-based randomised approach strengthens internal validity while preserving ecological cues, a recommended strategy for identifying causal effects in complex social settings (Atzmüller and Steiner, 2010). Incorporating RSS into the choice set reflects contemporary multimodality and enables safety and security perceptions to be compared across public and private options; this responds to evidence that perceived security is a salient driver of ride-hailing uptake, particularly off-peak (Clewlow and Mishra, 2017; Dias et al., 2017).

Incorporating a realistically priced RSS strengthens ecological validity. In our design, the RSS price was approximately five times the bus fare, and only 3.73% of participants selected it, consistent with strong price sensitivity that constrains substitution towards RSS when public transport remains markedly cheaper and time-competitive (Ward et al., 2021). This choice pattern aligns with external evidence showing context-dependent, often modest aggregate effects of ride-hailing on transit usage (significant bus ridership declines in many US metropolitan areas after TNC entry, but no statistically detectable association with ridership drops in several Latin American cities), indicating that market-level displacement can be limited where transit service and relative prices are favourable. Moreover, studies document that ride-hailing commonly increases vehicle miles travelled and emissions absent high pooling or electrification, suggesting that any perceived safety premium of RSS does not mechanically translate into broader efficiency gains (Sheldon and Dua, 2024). Our experiment mirrors this nuance by observing minimal displacement from bus to RSS at the offered relative prices. Together, the observed low RSS uptake at realistic relative prices and the small, imprecise shifts in bus uptake under enforcement cues reproduce key market regularities supporting ecological validity for policy-relevant counterfactuals.

4.2. Policy implications and recommendations

Before presenting practice-oriented recommendations, it is important to distinguish between findings supported by robust evidence and those that remain uncertain. Three findings are consistent across model specifications and carry direct policy implications: perceived route safety is a significant predictor of route choice ($OR = 1.866$, $p = 0.002$); individual fare-evasion history consistently predicts evasion behaviour; and night-time travel substantially increases hard-evasion propensity ($OR = 3.12\text{--}4.93$). By contrast, the effectiveness of inspection and fine cues in improving safety and security perceptions or reducing fare-evasion intentions remains empirically uncertain: observed effect sizes for enforcement contrasts were uniformly negligible, and these null results should be interpreted as contextually bounded rather than as a general refutation of deterrence-based enforcement. The four recommendations below reflect this distinction, moving from the limits of current enforcement tools to the positive levers identified by the data.

First, TAs/PTCs should not infer from these findings that informational enforcement cues are sufficient on their own to boost ridership through reassurance. Across all specifications, inspection and fine cues did not measurably improve perceived safety or security (RQ1), nor did they increase the attractiveness of the manipulated route through the perceptual pathway hypothesised in RQ2. The Fine condition was, if anything, associated with a lower likelihood of route selection than the Inspection condition ($OR = 0.456$, $p = 0.037$), suggesting that sanction-focused messaging may be perceived as coercive rather than reassuring (Fürst and Herold, 2018; Suquet, 2010). These results are consistent with deterrence research showing that behavioural effects of sanction severity and detection probability are context-dependent and may be attenuated when policy-relevant variations are modest rather than extreme (Chalfin and McCrary, 2017; Nagin, 2013). The dual-pathway framework introduced in Section 1 offers an explanation: fine-based cues operate through perceived sanction severity without generating immediate guardianship signals for observing passengers, whereas visible inspection presence may more directly activate the perceptual pathway. Therefore, policy should prioritise integrated measures that make guardianship credibly visible while improving the broader service environment, rather than relying on marginal adjustments to fine schedules alone.

Second, TAs/PTCs should actively manage safety perceptions as a co-equal performance indicator alongside evasion rates. The most robust and direct finding of this study is that perceived comparative route safety strongly predicts route choice ($OR = 1.866$, 95% CI [1.259, 2.764], $p = 0.002$): passengers who perceived the manipulated route as comparatively safer were significantly more likely to select it. Yet enforcement cues did not reliably shift these perceptions in the present experiment. This decoupling between enforcement intensity and safety perception implies that perceived safety is shaped by factors beyond formal deterrence, including environmental quality, service reliability, waiting conditions, and lighting. As a result, TAs/PTCs should track perceived safety and security as structured performance indicators and address their underlying determinants directly, rather than assuming that enforcement outputs will automatically translate into passenger reassurance (Clarke et al., 2010).

Third, night-time travel warrants dedicated and differentiated policy intervention. Night-time conditions were associated with substantially higher hard-evasion propensity across all specifications ($OR = 3.12\text{--}4.93$), corresponding to a standardised effect of approximately $d = 0.63\text{--}0.88$. Critically, this effect operates independently of enforcement condition: RQ3 found no evidence that inspection or fine cues differentially improved safety or security perceptions at night relative to daytime travel. Therefore, the behavioural salience of night-time context does not appear to be addressed by marginal adjustments to enforcement intensity. TAs/PTCs should instead package night-specific interventions (including predictable service headways, illuminated and monitored waiting environments, staffed or supervised interchange hubs, and clearly signalled request-stop protocols) to strengthen capable guardianship after dark (Ceccato and Loukaitou-Sideris, 2020; Cohen and Felson, 1979).

Fourth, TAs/PTCs should embed procedural justice into inspection protocols, prioritising the quality of enforcement interactions over their intensity. In this study, neither fine severity nor inspection frequency was found to significantly improve perceived safety or compliance intentions. This suggests that the effectiveness of fare enforcement may depend not only on how often inspections are carried out or how severe penalties are, but also on the way inspections are performed and experienced by passengers. The robust predictive role of individual fare-evasion history across specifications suggests that evasion behaviour is partly shaped by persistent

habits, prior experiences, and perceived norms; namely factors that marginal enforcement cues are unlikely to disrupt. Positioning inspectors as visible, respectful guardians who help, clear information, and procedurally fair treatment is more likely to strengthen perceived legitimacy and capable guardianship than simply escalating sanctions (Barabino and Salis, 2019; Mazerolle et al., 2013). This approach is grounded in routine activity theory's emphasis on the quality of capable guardianship as a mechanism for reducing opportunities for deviant behaviour and improving perceived order in public environments (Cohen and Felson, 1979; Hollis et al., 2013).

4.3. Limitations and future research

We identify some limitations that warrant caution and need of future studies.

First, the intention–behaviour gap to vignette-based designs is a challenge in behavioural research. Although the scenario was designed to maximise realism through anonymity and context-specific framing, stated intentions remain imperfect proxies for actual behaviour (Sheeran, 2002; Sheeran and Webb, 2016). Nevertheless, intentions constitute the strongest proximal predictor of behaviour in meta-analytic evidence (Armitage and Conner, 2001). In addition, the predictive validity of vignette-elicited responses has been supported by direct comparisons with natural experiments across related applied settings (Antinyan et al., 2021; Struminskaya et al., 2021). Design features aimed at reducing socially desirable responses, including anonymity and non-moralising framing, were systematically implemented (Aguinis and Bradley, 2014; Eifler and Petzold, 2019). As a result, the intention estimates should be treated as conservative, policy-relevant proxies for behavioural propensity, and recommend that future research triangulates these findings against administrative inspection data or unobtrusive behavioural measures.

Second, external validity is constrained by the sample composition, which consists overwhelmingly of young university students. This may limit the generalisability of the findings to the broader population of public transport users. However, this group is particularly relevant for analysing fare evasion, as prior studies consistently identify younger individuals and students (including university students) among those more likely to evade fares, as well as representing a substantial share of public transport users (e.g., Allen et al., 2019; Barabino et al., 2015; Barabino and Salis, 2020; Buccioli et al., 2013). At the same time, fare evasion behaviour is known to vary significantly across socioeconomic contexts and user groups. Evidence from different settings shows that higher evasion rates are often associated with more vulnerable populations and low-income areas (e.g., Munizaga et al., 2020). This suggests that the patterns observed in this study may not fully capture behaviours present in more socioeconomically diverse contexts. Moreover, given the heterogeneity of fare evasion behaviour across user segments, focusing on a relatively homogeneous group allows for a more controlled analysis of behavioural mechanisms (Delbosc and Currie, 2016). Nevertheless, caution is required when generalising the results. Future research should therefore include more socioeconomically diverse user groups to improve the external validity of the findings (Guzman et al., 2021; Munizaga et al., 2020).

Third, methodologically, vignettes cannot reproduce all in-field contingencies (crowding, lighting, peer effects). A pragmatic next step is to pair vignettes with field or quasi-experimental tests that vary inspection presence across routes and time bands and evaluate causal impacts with operational data (Ayal et al., 2021; Clarke et al., 2010; Troncoso and de Grange, 2017).

Fourth is that SaRI and SeRI were treated as observed composite indices rather than latent variables. This choice preserved parsimony and estimation feasibility in a multi-equation GSEM with a limited sample size, but it necessarily treats item-level measurement error only indirectly. Although CFA-based factor-score robustness checks produced substantively unchanged conclusions, future studies should replicate the design with larger samples and estimate SaRI and SeRI as latent constructs within an integrated SEM framework. This would allow the measurement model, the covariance between safety and security perceptions, and the downstream structural pathways to be estimated jointly.

Fifth, although total travel time was kept approximately constant across alternatives, differences in walking, waiting and in-vehicle time might have affected felt generalised cost, as walking and waiting time are typically perceived as more burdensome than in-vehicle time. This may have influenced route choices independently of the experimental treatments. More research is needed to address this issue by explicitly controlling for perceived generalised cost components.

Finally, several extensions could deepen inference and policy relevance. Future analyses could estimate moderated effects by gender, prior victimisation, and habitual RSS use to identify audiences for targeted night-time interventions (Ceccato and Loukaitou-Sideris, 2020; Loukaitou-Sideris, 2014). They could also examine whether crime prevention through environmental design-informed environmental upgrades (lighting uniformity, sightlines, staffed presence) and operational policies (shorter headways to reduce waiting), combined with inspection presence shift both perceived and actual safety (Newman, 1996; Saxe, 2024). Moreover, integrating operator cost data and enforcement logs could allow an economic calibration of deterrence, quantifying the marginal effectiveness of inspection probability versus fine severity along the probability–severity frontier in public enforcement theory (Polinsky and Shavell, 2000; Troncoso and de Grange, 2017). Finally, future research could incorporate spatio-temporal risk models that distinguish static (stops/stations) from moving (on-board) crime events when modelling night-time exposure and guardianship (Newton, 2004).

Ethical approval

This study received ethical approval from the Ethical Committee of ESSCA School of Management (approval number: 10/2024).

CRedit authorship contribution statement

Benedetto Barabino: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation,

Conceptualization. **Martina Carra**: Writing – review & editing, Conceptualization. **Jérémy Celse**: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Roberto Ventura**: Writing – review & editing.

Funding

No funding was requested for this study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

During the preparation of this work, the authors used ChatGPT 5.5 (OpenAI) to improve English language and readability. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.tra.2026.105244>.

Data availability

Data will be made available on request.

References

- Aguinis, H., Bradley, K.J., 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organ. Res. Methods* 17 (4), 351–371. <https://doi.org/10.1177/1094428114547952>.
- Alade, T., Edelenbos, J., Gianoli, A., 2020. A sustainable approach to innovation adoption in light-rail transport. *Sustainability* 12 (3), 1262. <https://doi.org/10.3390/su12031262>.
- Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2018. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behav. Soc.* 13, 88–104. <https://doi.org/10.1016/j.tbs.2018.06.002>.
- Allen, J., Muñoz, J.C., de Ortúzar, J., 2019. On evasion behaviour in public transport: dissatisfaction or contagion? *Transp. Res. A Policy Pract.* 130, 626–651. <https://doi.org/10.1016/j.tra.2019.10.005>.
- Antinyan, A., Bassetti, T., Corazzini, L., Pavesi, F., 2021. Trust in the health system and COVID-19 treatment. *Front. Psychol.* 12. <https://doi.org/10.3389/fpsyg.2021.643758>.
- Armitage, C.J., Conner, M., 2001. Efficacy of the theory of planned behaviour: a meta-analytic review. *Br. J. Soc. Psychol.* 40 (4), 471–499. <https://doi.org/10.1348/01446660164939>.
- Assaf, C., Van den Broeck, P., 2022. The quest for a fairer formula. How re-institutionalisation begins with neo-illegal transport communities at the fringe of social innovation in Brussels. *Eur. J. Spatial Develop.* 19 (4), 72–98.
- Atzmüller, C., Steiner, P.M., 2010. Experimental vignette studies in survey research. *Methodology* 6 (3), 128–138. <https://doi.org/10.1027/1614-2241/a000014>.
- Ayal, S., Celse, J., Hochman, G., 2021. Crafting messages to fight dishonesty: a field investigation of the effects of social norms and watching eye cues on fare evasion. *Organ. Behav. Hum. Decis. Process.* 166, 9–19. <https://doi.org/10.1016/j.obhdp.2019.10.003>.
- Barabino, B., Bonera, M., Maternini, G., Porcu, F., Ventura, R., 2024a. Refining a crash risk framework for urban bus safety assessment: evidence from Sardinia (Italy). *Reliab. Eng. Syst. Saf.* 245, 110003. <https://doi.org/10.1016/j.res.2024.110003>.
- Barabino, B., Carra, M., Currie, G., 2024b. Fare inspection in proof-of-payment transit networks: a review. *J. Public Transp.* 26, 100101. <https://doi.org/10.1016/j.jpubtr.2024.100101>.
- Barabino, B., Lai, C., Olivo, A., 2020. Fare evasion in public transport systems: a review of the literature. *Public Transp.* 12 (1), 27–88. <https://doi.org/10.1007/s12469-019-00225-w>.
- Barabino, B., Salis, S., 2019. Moving towards a more accurate level of inspection against fare evasion in proof-of-payment transit systems. *Netw. Spat. Econ.* 19 (4), 1319–1346. <https://doi.org/10.1007/s11067-019-09468-3>.
- Barabino, B., Salis, S., 2020. Do students, workers, and unemployed passengers respond differently to the intention to evade fares? An empirical research. *Transp. Res. Interdiscip. Perspect.* 7, 100215. <https://doi.org/10.1016/j.trip.2020.100215>.
- Barabino, B., Salis, S., 2023. Segmenting fare-evaders by tandem clustering and logistic regression models. *Public Transp.* 15 (1), 61–96. <https://doi.org/10.1007/s12469-022-00297-1>.
- Barabino, B., Salis, S., Useli, B., 2015. What are the determinants in making people free riders in proof-of-payment transit systems? Evidence from Italy. *Transp. Res. A Policy Pract.* 80, 184–196. <https://doi.org/10.1016/j.tra.2015.07.017>.
- Becker, G.S., 1968. Crime and punishment: an economic approach. *J. Polit. Econ.* 76 (2), 169. <https://doi.org/10.1086/259394>.
- Berryessa, C.M., Caplan, J.M., 2020. Cognitive and affective processing of risk information: a survey experiment on risk-based decision-making related to crime and public safety. *Front. Psychol.* 11. <https://doi.org/10.3389/fpsyg.2020.02222>.
- Beylveeld, D., 1979a. Deterrence research as a basis for deterrence policies. *Howard J. Crim. Just.* 18 (3), 135–149. <https://doi.org/10.1111/j.1468-2311.1979.tb00388.x>.
- Beylveeld, D., 1979b. Identifying, explaining and predicting deterrence. *British J. Criminol.* 19 (3), 205–224. <https://doi.org/10.1093/oxfordjournals.bjc.a047002>.
- Bijlveeld, C., 2007. Fare dodging and the strong arm of the law: an experimental evaluation of two different penalty schemes for fare evasion. *J. Exp. Criminol.* 3 (2), 183–199. <https://doi.org/10.1007/s11292-007-9033-1>.
- Bonfanti, G., Wagenknecht, T., 2010. Human factors reduce aggression and fare evasion. *Public Transp. Int.* 59 (1), 28–32.

- Brodeur, A., Cook, N.M., Hartley, J.S., Heyes, A., 2024. Do Preregistration and preanalysis plans reduce p-hacking and publication bias? Evidence from 15,992 test statistics and suggestions for improvement. *J. Polit. Econ. Microecon.* 2 (3), 527–561. <https://doi.org/10.1086/730455>.
- Buccioli, A., Landini, F., Piovesan, M., 2013. Unethical behavior in the field: demographic characteristics and beliefs of the cheater. *J. Econ. Behav. Organ.* 93, 248–257. <https://doi.org/10.1016/j.jebo.2013.03.018>.
- Ceccato, V., Gaudet, N., Graf, G., 2022. Crime and safety in transit environments: a systematic review of the English and the French literature, 1970–2020. *Public Transp.* 14 (1), 105–153. <https://doi.org/10.1007/s12469-021-00265-1>.
- Ceccato, V., Loukaitou-Sideris, A., 2020. Transit crime and sexual violence in cities (V. Ceccato & A. Loukaitou-Sideris, Eds.). Routledge. 10.4324/9780429290244.
- Ceder, A., 2019. *Public transit planning and operation : modeling, practice and behavior*. CRC Press.
- Celse, J., Grolleau, G., 2023. Fare evasion and information provision: what information should be provided to reduce fare-evasion? *Transp. Policy* 138, 119–128. <https://doi.org/10.1016/j.tranpol.2023.05.008>.
- Chalfin, A., McCrary, J., 2017. Criminal deterrence: a review of the literature. *J. Econ. Lit.* 55 (1), 5–48. <https://doi.org/10.1257/jel.20141147>.
- Chen, H., Cohen, P., Chen, S., 2010. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun. Stat. Simul. Comput.* 39 (4), 860–864. <https://doi.org/10.1080/03610911003650383>.
- Clarke, R.V., Contre, S., Petrossian, G., 2010. Deterrence and fare evasion: results of a natural experiment. *Secur. J.* 23 (1), 5–17. <https://doi.org/10.1057/sj.2009.15>.
- Clewlow, R.R., Mishra, G.S., 2017. Disruptive transportation: the adoption, utilization, and impacts of ride-hailing in the United States.
- Cohen J., 1988. Statistical power analysis for the behavioural science (2nd Edition). In Routledge (Ed.), *Statistical Power Analysis for the Behavioural Science* (2nd Edition).
- Cohen, L.E., Felson, M., 1979. Social change and crime rate trends: a routine activity approach. *Am. Sociol. Assoc.* 44 (4), 588–608.
- Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16 (3), 297–334. <https://doi.org/10.1007/BF02310555>.
- Currie, G., Delbosc, A., Mahmoud, S., 2013. Factors influencing young peoples' perceptions of personal safety on public transport. *J. Public Transp.* 16 (1), 1–19. <https://doi.org/10.5038/2375-0901.16.1.1>.
- Delbosc, A., Currie, G., 2012. Modelling the causes and impacts of personal safety perceptions on public transport ridership. *Transp. Policy* 24, 302–309. <https://doi.org/10.1016/j.tranpol.2012.09.009>.
- Delbosc, A., Currie, G., 2016. Cluster analysis of fare evasion behaviours in Melbourne, Australia. *Transp. Policy* 50, 29–36. <https://doi.org/10.1016/j.tranpol.2016.05.015>.
- Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., Kaiser, S., 2012. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *J. Acad. Mark. Sci.* 40 (3), 434–449. <https://doi.org/10.1007/s11747-011-0300-3>.
- Dias, F.F., Lavieri, P.S., Garikapati, V.M., Astroza, S., Pendyala, R.M., Bhat, C.R., 2017. A behavioral choice model of the use of car-sharing and ride-sourcing services. *Transportation* 44 (6), 1307–1323. <https://doi.org/10.1007/s11116-017-9797-8>.
- Eck, J.E., Clarke, R.V., 2019. Situational crime prevention: theory, practice and evidence. In: Krohn, M.D., Hendrix, N., Penly Hall, G., Lizotte, A.J. (Eds.), *Handbook on Crime and Deviance*. Springer International Publishing, pp. 355–376. https://doi.org/10.1007/978-3-030-20779-3_18.
- Eifler, S., Petzold, K., 2019. Validity aspects of vignette experiments: expected “what-if” differences between reports of behavioral intentions and actual behavior. In *Experimental Methods in Survey Research* (pp. 393–416). Wiley. 10.1002/9781119083771.ch20.
- Elvik, R., 2019. Risk of non-collision injuries to public transport passengers: synthesis of evidence from eleven studies. *J. Transp. Health* 13, 128–136. <https://doi.org/10.1016/j.jth.2019.03.017>.
- Escalona, P., Brotcorne, L., Fortz, B., Ramirez, M., 2024. Fare inspection patrolling under in-station selective inspection policy. *Ann. Oper. Res.* 332 (1–3), 191–212. <https://doi.org/10.1007/s10479-023-05670-2>.
- Escalona, P., Brotcorne, L., Fortz, B., Wolf, N., 2025. Spot-fare inspection in urban bus transportation systems: strategy and unpredictability under a Stackelberg game approach. *Public Transp.* <https://doi.org/10.1007/s12469-025-00398-7>.
- Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods*. <https://doi.org/10.3758/BF03193146>.
- Finch, J., 1987. The vignette technique in survey research. *Sociology* 21 (1), 105–114. <https://doi.org/10.1177/0038038587021001008>.
- Freiria, S., Sousa, N., 2025. Determinants of fare evasion in urban bus lines: case study of a large database considering spatial components. *Urban Sci.* 9 (6), 231. <https://doi.org/10.3390/urbansci9060231>.
- Freiria, S., Sousa, N., 2026. A spatial statistics methodology for inspector allocation against fare evasion. *ISPRS Int. J. Geo Inf.* 15 (2), 53. <https://doi.org/10.3390/ijgi15020053>.
- Fruin, J., Smith, N.E., Huang, H.F., Zegeer, C.V., 1994. Recommendations for reducing noncollision bus passenger injuries. *Transportation Record* 1433, 41–50.
- Fürst, E.W.M., Herold, D.M., 2018. Fare evasion and ticket forgery in public transport: insights from Germany Austria and Switzerland. *Societies* 8 (4), 98. <https://doi.org/10.3390/soc8040098>.
- González, F., Busco, C., Codocedo, K., 2019. Fare evasion in public transport: grouping transantiago users' behavior. *Sustainability* 11 (23). <https://doi.org/10.3390/su11236543>.
- Guzman, L.A., Arellana, J., Camargo, J.P., 2021. 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.* 151, 140–153. <https://doi.org/10.1016/j.tra.2021.07.009>.
- Hauber, A.R., 1993. Fare evasion in a European perspective. *Studies on Crime and Crime Prevention* 2, 122–141.
- Hirsch, A.V., Bottoms, A., Burney, E., Wikström, P.-O., 1999. Criminal deterrence and sentence severity: an analysis of recent research. Hart Publishing, Oxford, pp. 2–5. <https://doi.org/10.1017/CBO9781107415324.004>.
- Hollis, M.E., Felson, M., Welsh, B.C., 2013. The capable guardian in routine activities theory: a theoretical and conceptual reappraisal. *Crime Prev. Community Saf.* 15 (1), 65–79. <https://doi.org/10.1057/cpcs.2012.14>.
- Hughes, F., Huby, M., 2004. The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review - an International Journal of Applied* 11, 36–51. <https://doi.org/10.1921/17466105.11.1.36>.
- Johnson, L.T., Patterson, E.J., 2022. The policing of subway fare evasion in postindustrial Los Angeles. *Punishment Soc.* 24 (3), 457–476. <https://doi.org/10.1177/1462474521992115>.
- Kaplan, S., Bekhor, S., Shiftan, Y., 2012. Web-based survey design for unravelling semi-compensatory choice in transport and urban planning. *Transp. Plan. Technol.* 35 (2), 121–143. <https://doi.org/10.1080/03081060.2012.651874>.
- Keuchel, S., Swertz, C., 2020. Effects of ticket inspection levels: a quasi-experimental approach. *Transp. Res. Procedia* 48, 1824–1834. <https://doi.org/10.1016/j.trpro.2020.08.217>.
- Keuschnigg, M., Wolbring, T., 2015. Experimente in den Sozialwissenschaften. 22. Sonderband der Sozialen Welt (M. Keuschnigg & T. Wolbring, Eds.; Vol. 22). Nomos. 10.5771/9783845260433-1.
- Killias, M., Scheidegger, D., Nordenson, P., 2009. The effects of increasing the certainty of punishment. *Eur. J. Criminol.* 6 (5), 387–400. <https://doi.org/10.1177/1477370809337881>.
- Lakens, D., 2017. Equivalence tests. *Soc. Psychol. Personal. Sci.* 8 (4), 355–362. <https://doi.org/10.1177/1948550617697177>.
- Logg, J.M., Dorison, C.A., 2021. Pre-registration: weighing costs and benefits for researchers. *Organ. Behav. Hum. Decis. Process.* 167, 18–27. <https://doi.org/10.1016/j.obhdp.2021.05.006>.
- Loukaitou-Sideris, A., 2014. Fear and safety in transit environments from the women's perspective. *Secur. J.* 27 (2), 242–256. <https://doi.org/10.1057/sj.2014.9>.
- Louviere, J.J., Hensher, D.A., Swait, J.D., Adamowicz, W., 2000. Transportation case studies. In: Louviere, J.J., Hensher, D.A., Swait, J.D. (Eds.), *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, pp. 298–328. <https://doi.org/10.1017/CBO9780511753831.011>.
- Macé, E., 1997. Service public et banlieues populaires : une coproduction de l'insécurité. Le cas du réseau bus de la RATP. *Sociologie Du Travail* 39 (4), 473–498. <https://doi.org/10.3406/sotra.1997.2329>.
- MacKinnon, D.P., 2012. *Introduction to Statistical Mediation Analysis*. Routledge. <https://doi.org/10.4324/9780203809556>.

- Mauri, R.A., Cooney, N.A., Prowe, G.J., 1984. Transit security: a description of problems and countermeasures.
- Mazerolle, L., Bennett, S., Davis, J., Sargeant, E., Manning, M., 2013. Legitimacy in policing: a systematic review. *Campbell Syst. Rev.* 9 (1). <https://doi.org/10.4073/csr.2013.1>.
- Munizaga, M.A., Gschwendner, A., Gallegos, N., 2020. Fare evasion correction for smartcard-based origin-destination matrices. *Transp. Res. A Policy Pract.* 141, 307–322. <https://doi.org/10.1016/j.tra.2020.09.008>.
- Muthén, B., 1984. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 49 (1), 115–132. <https://doi.org/10.1007/BF02294210>.
- Nagin, D.S., 2013. Deterrence in the twenty-first century. *Crime Justice* 42, 199–263. <https://doi.org/10.1086/670398>.
- Nakanishi, Y.J., Fleming, W.C., 2011. Practices to protect bus operators from passenger assault. National Academies Press. 10.17226/14609.
- Newman, O., 1996. Creating defensible space. U.S. Department of Housing and Urban Development. Office of Policy Development and Research.
- Newton, A., 2004. Crime on public transport: 'static' and 'non-static' (moving) crime events. *West. Criminol. Rev.* 5 (3), 25–42.
- Norman, G., 2010. Likert scales, levels of measurement and the "laws" of statistics. *Adv. Health Sci. Educ.* 15 (5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>.
- Nunnally, J.C., Bernstein, I.H., 1994. *Psychometric theory*. McGraw-Hill.
- Ortuzar, J. de Dios, Willumsen, L.G., 2011. *Modelling Transport*. Wiley [Imprint] John Wiley & Sons, Incorporated.
- Polinsky, A.M., Shavell, S., 2000. The economic theory of public enforcement of law. *J. Econ. Lit.* 38 (1), 45–76. <https://doi.org/10.1257/jel.38.1.45>.
- Porath, K., Galilea, P., 2020. Temporal analysis of fare evasion in transantiago: a socio-political view. *Res. Transp. Econ.* 83, 100958. <https://doi.org/10.1016/j.retrec.2020.100958>.
- Preacher, K.J., Hayes, A.F., 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* 40 (3), 879–891. <https://doi.org/10.3758/BRM.40.3.879>.
- Rabe-Hesketh, S., Skrondal, A., 2012. *Multilevel and longitudinal modeling using stata (third edition): I*. Stata Press, College Station, TX.
- Raykov, T., Marcoulides, G.A., 2011. *Introduction to psychometric theory*. Routledge.
- Reddy, A.V., Kuhls, J., Lu, A., 2011. Measuring and controlling subway fare evasion: improving safety and security at New York City Transit Authority. *Transp. Res. Rec.* 2216, 85–99.
- Saxe, J.P., 2024. Crime prevention through environmental design (CPTED) for public transit stations.
- Shaheen, S., Cohen, A., 2019. Shared ride services in North America: definitions, impacts, and the future of pooling. *Transp. Rev.* 39 (4), 427–442. <https://doi.org/10.1080/01441647.2018.1497728>.
- Sheeran, P., 2002. Intention—behavior relations: a conceptual and empirical review. *Eur. Rev. Soc. Psychol.* 12 (1), 1–36. <https://doi.org/10.1080/14792772143000003>.
- Sheeran, P., Webb, T.L., 2016. The intention—behavior gap. *Soc. Pers. Psychol. Compass* 10 (9), 503–518. <https://doi.org/10.1111/spc3.12265>.
- Sheldon, T.L., Dua, R., 2024. Impacts of ride-hailing on energy and the environment: a systematic review. *Environ. Res. Lett.* 19 (4), 043004. <https://doi.org/10.1088/1748-9326/ad3285>.
- Skogan, W.G., 1992. *Disorder and decline : crime and the spiral of decay in American neighborhoods*. University of California Press.
- Smith, M., Cornish, D., 2006. *Secure and tranquil travel : preventing crime and disorder on public transport*. Jill Dando Institute of Crime, Science University College London.
- Steiner, P., Atzmüller, C., Su, D., 2017. Designing valid and reliable vignette experiments for survey research: a case study on the fair gender income gap. *J. Methods Meas. Soc. Sci.* 7 (2), 52–94.
- Sterner, A., Sheng, S., 2013. The effect of social stigma on fare evasion in Stockholm's public transport. *J. Transp. Literature* 7, scielo.
- Stoet, G., 2010. PSYTOOLKIT: a software package for programming psychological experiments using Linux. *Behav. Res. Methods* 42 (4), 1096–1104. <https://doi.org/10.3758/BRM.42.4.1096>.
- Stoet, G., 2016. psytoolkit: a novel web-based method for running online questionnaires and reaction-time experiments. *Teach. Psychol.* 44 (1), 24–31. <https://doi.org/10.1177/0098628316677643>.
- Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., Dolmans, R., 2021. Sharing data collected with smartphone sensors. *Public Opin. Q.* 85, 423–462. <https://api.semanticscholar.org/CorpusID:238248602>.
- Sullivan, G.M., Artino, A.R., 2013. Analyzing and interpreting data from likert-type scales. *J. Grad. Med. Educ.* 5 (4), 541–542. <https://doi.org/10.4300/JGME-5-4-18>.
- Sundling, C., Ceccato, V., Giori, G., 2024. Disability, victimisation, and safety in train travel. *Transp. Res. Interdiscip. Perspect.* 25, 101131. <https://doi.org/10.1016/j.trip.2024.101131>.
- Suquet, J., 2010. Drawing the line: how inspectors enact deviant behaviors. *J. Serv. Mark.* 24 (6), 468–475. <https://doi.org/10.1108/08876041011072582>.
- Taber, K.S., 2018. The use of cronbach's alpha when developing and reporting research instruments in science education. *Res. Sci. Educ.* 48 (6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>.
- Too, L., Earl, G., 2010. Public transport service quality and sustainable development: a community stakeholder perspective. *Sustain. Dev.* 18 (1), 51–61. <https://doi.org/10.1002/sd.412>.
- Tourangeau, R., Yan, T., 2007. Sensitive questions in surveys. *Psychol. Bull.* 133 (5), 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>.
- Troncoso, R., de Grange, L., 2017. Fare evasion in public transport: a time series approach. *Transp. Res. A Policy Pract.* 100, 311–318. <https://doi.org/10.1016/j.tra.2017.04.029>.
- van Dijk, J.J.M., Junger-Tas, J., 1988. Trends in crime prevention in the Netherlands. In: Hope, T., Shaw, M. (Eds.), *Communities and Crime Reduction*. Home Office Research and Planning Unit, pp. 260–277.
- Ward, J.W., Michalek, J.J., Samaras, C., Azevedo, I.L., Henao, A., Rames, C., Wenzel, T., 2021. The impact of Uber and Lyft on vehicle ownership, fuel economy, and transit across U.S. cities. *Iscience* 24 (1), 101933. <https://doi.org/10.1016/j.isci.2020.101933>.
- Wilson, J., Kelling, G., 1982. Broken windows: the police and neighborhood safety. *The Atlantic Monthly* 29–38.
- Zhong, C.-B., Bohns, V.K., Gino, F., 2010. Good lamps are the best police: darkness increases dishonesty and self-interested behavior. *Psychol. Sci.* 21 (3), 311–314. <https://doi.org/10.1177/0956797609360754>.