Cities 178 (2026) 107350

Contents lists available at ScienceDirect

Cities

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

Exploring cargo bike viability for last-mile delivery in European capitals: A mass-participation multi-city Multi-Criteria Satisfaction Analysis

River Huang

Laboratory for Energy Systems Analysis, PSI Centers for Nuclear Engineering & Sciences and Energy & Environmental Sciences, Forschungsstrasse 111, Villigen PSI, 5232, Switzerland

ARTICLE INFO

Keywords:
Mass-participation
Cargo bike
Logistics
MUSA
MCDA
Capital cities

ABSTRACT

The rapid growth of e-commerce has intensified urban freight challenges and increased interest in cargo bikes as a potentially more sustainable last-mile delivery option. However, public acceptance of cargo bike deliveries remains underexplored. However, comparative evidence on citizens' acceptance of such services across cities remains limited. This study conducts a mass-participation survey across six European capital cities to evaluate citizens' acceptance of a hypothetical cargo bike delivery service ("EcoDeliver"). We apply a multi-city MUlticriteria Satisfaction Analysis (MUSA) model to analyze overall acceptance and criterion-level evaluations across environmental impact, noise, traffic congestion, safety, and shipping cost. Approximately 2500 citizens assessed the projected performance of the proposed service scenario. The results indicate generally positive evaluations of the hypothetical service, with environmental impact, noise reduction, and congestion reduction receiving the most favorable responses overall. At the same time, the relative importance of the evaluation criteria and the shape of the acceptance functions vary across cities and demographic groups, revealing differences in how citizens prioritize key attributes of urban logistics services. These results provide a more differentiated understanding of public preferences and show that similar overall acceptance levels may be associated with distinct local priority structures. Methodologically, the study illustrates how MUSA can be used to structure and compare large-scale citizen evaluations in participatory urban logistics research.

1. Introduction

The booming growth of e-commerce and on-demand deliveries has profoundly impacted urban logistics, resulting in smaller, more frequent shipments and increasing last-mile delivery traffic (Bell, 2021). This trend exacerbates urban challenges such as air pollution, noise, congestion, and road safety issues. Cities house roughly 70% of Europe's population, and about 23% of EU transport greenhouse gas emissions originate from urban areas (European Commission, 2025). In response, ambitious targets like a 90% reduction in transport emissions by 2050 have been set under the European Green Deal (Hainsch et al., 2022). Policy frameworks, including Sustainable Urban Mobility Plans (SUMPs) and Sustainable Urban Logistics Plans (SULPs), alongside the EU's Urban Mobility Framework, increasingly advocate for sustainable solutions and public engagement in mobility (Ambrosino, 2014; Commission et al., 2013; European Commission, 2021).

Among the options proposed for more sustainable last-mile delivery, cargo bikes have received growing attention. Cargo bikes, particularly electric cargo bikes, produce no tailpipe emissions and require considerably less road space than conventional delivery vans, which may help reduce both environmental impacts and traffic pressure (Gunes, 2021). Their use has expanded in several European markets, including Germany, France, Belgium, and the Netherlands. Empirical trials and pilot applications suggest that cargo bikes can perform a meaningful share of urban deliveries while reducing emissions and noise (Michalakopoulou et al., 2025). At the same time, recent reviews indicate that their broader deployment depends on suitable operational designs and supportive urban conditions, such as cycling infrastructure, consolidation facilities, and micro-hubs (Chatziioannou et al., 2024).

Despite these promising developments, a critical gap remains in understanding the perspective of the general public, who are the end consumers and citizens affected by urban freight. Most research on cargo bikes has focused on operational feasibility, cost, and environmental performance, or on the business and regulatory aspects of implementation. Far fewer studies have examined public acceptability or perceptions of cargo bike deliveries. Citizens' attitudes are crucial because they can influence the success of sustainable logistics initiatives. If residents perceive a new delivery mode as beneficial (for example, cleaner air and quieter streets), they are more likely to support or utilize it; if not, even the best-planned green logistics measures may

E-mail address: river.huang@psi.ch.

https://doi.org/10.1016/j.cities.2026.107350
Received 22 January 2026; Received in revised form 11 May 2026; Accepted 30 June 2026

face resistance. Recent work has started to address this gap. For instance, França et al. (2025) found that citizens' perceptions of urban freight problems strongly shape their support for freight initiatives. In a comparison of a Global South and a European city, safety concerns for pedestrians and cyclists were paramount in the former, whereas environmental pollution issues were more salient in the latter (Parajuli & Pojani, 2018). This exemplifies how local context influences what the public values in urban freight policy. Overall, there is an important call for policy-makers to capture citizen opinions on last-mile delivery solutions and incorporate them into planning. Public participation in urban logistics planning remains challenging, often limited by resource constraints and insufficient tools for engaging laypeople. Traditional approaches reach relatively small numbers of people and may not adequately inform participants about complex logistics issues. Thus, innovative, inclusive methods are needed to gauge public preferences at scale and inform more citizen-centric urban freight policies.

This paper addresses that need through a mass-participation study of citizen evaluations of a hypothetical cargo-bike last-mile delivery service in six European capital cities. More than 2500 residents assessed a common service scenario, allowing direct comparison across cities under a consistent framing. The study has three objectives: (1) to assess overall public acceptance of the presented cargo-bike delivery scenario, (2) to identify the relative importance of key service attributes, namely environmental impact, noise, traffic congestion, safety, and shipping cost, and (3) to examine similarities and differences in preference structures across cities and demographic groups. To do so, we combine a standardized survey scenario with MUlticriteria Satisfaction Analysis (MUSA) (Grigoroudis & Siskos, 2002). This allows public evaluations to be analyzed not only in terms of overall acceptance, but also in terms of criterion importance and the shape of the corresponding acceptance value functions, thereby providing a structured account of how citizens evaluate a multi-attribute logistics service. In this way, the study extends existing freight-perception research by enabling harmonized multi-city comparison and by identifying the criterion-specific patterns that underlie overall acceptance. The resulting analysis provides policy-relevant insights into both shared and context-specific public priorities in sustainable urban logistics.

The remainder of the paper is organized as follows. Section 2 reviews the literature on public participation in sustainable urban logistics and on cargo-bike adoption. Section 3 presents the methodological framework, including the MUSA model and the survey design. Section 4 describes the survey implementation and sample. Section 5 reports the empirical results. Section 6 discusses the findings and their implications for urban logistics planning. The final section concludes with the main findings, limitations, and directions for future research.

2. Literature review

Public participation is increasingly regarded as an essential component of sustainable urban mobility and logistics planning. In the European context, this orientation is reflected in policy frameworks such as Sustainable Urban Mobility Plans (SUMPs) and Sustainable Urban Logistics Plans (SULPs), which promote the inclusion of citizen perspectives alongside those of public authorities, businesses, and other organized stakeholders (Roman & Fellnhofer, 2022). In practice, however, meaningful participation in urban freight remains difficult to achieve. Existing approaches such as workshops, focus groups, and living labs can generate rich qualitative insights, but they are often resource-intensive, limited in scale, and difficult to compare across cities (Maltese et al., 2023; Mladenovič et al., 2022). In addition, several studies note that citizens may have only limited prior knowledge of urban freight systems, which complicates the elicitation of informed judgments on logistics interventions (Kiba-Janiak et al., 2021). As a result, urban logistics planning requires participatory approaches that are broad enough to capture heterogeneous public views, while remaining structured enough to yield analytically useful results.

This need is particularly important because public evaluations of urban freight are highly context-dependent. What citizens perceive as the main freight-related problems, and which improvements they value most, may differ substantially across places and social groups. Foltỳnová et al. (2020) showed that understandings of sustainable urban mobility can diverge across stakeholder categories, while França et al. (2025) found that perceived freight externalities shape support for urban freight initiatives. Related evidence suggests that local context also affects which issues become most salient, for example safety in one setting and environmental quality in another (Parajuli & Pojani, 2018). Taken together, this literature indicates that citizen evaluations should not be treated as homogeneous or readily transferable across cities. Yet, freight research still offers relatively little evidence based on a common analytical framework that allows direct comparison of public evaluations across multiple urban contexts.

Cargo bikes have emerged as one of the most visible strategies for improving the sustainability of last-mile logistics. A substantial part of the literature has examined their operational feasibility, environmental performance, routing efficiency, and business potential. This work generally shows that cargo bikes can perform a meaningful share of urban deliveries, particularly in dense areas and when combined with micro-hubs or other consolidation arrangements (Lenz & Riehle, 2013; Llorca & Moeckel, 2021; Philips et al., 2025; Robichet et al., 2022; Rudolph & Gruber, 2017). Reviews further indicate that successful deployment depends not only on vehicle performance, but also on infrastructure, regulation, and urban design conditions (Chatziioannou et al., 2024). This point is reinforced by the recent study of Chatziioannou et al. (2026), which highlights the role of dedicated bicycle infrastructure, street lighting, and broader pro-cycling policy environments in supporting cargo-bike safety and ease of movement. These findings suggest that cargo-bike viability is not purely a technical or operational issue, but is shaped by a wider urban environment that influences both actual use and public perception.

By contrast, the public-acceptance dimension remains much less developed. Existing studies have begun to examine how consumers and citizens view alternative last-mile delivery concepts, including their willingness to accept greener delivery options or to pay more for them Caspersen and Navrud (2021) and Hagen and Scheel-Kopeinig (2021). While these studies are valuable, three limitations remain. First, many focus on consumers as service users, whereas urban freight interventions also affect non-users as residents, pedestrians, cyclists, and neighbors. Second, most available evidence comes from single-city or otherwise context-specific studies, which limits cross-city comparison. Third, the dominant empirical approaches tend to yield descriptive summaries or isolated attribute effects rather than a structured view of how multiple service attributes jointly shape overall acceptance and where the most relevant intervention levers lie. Consequently, the literature still provides limited guidance on how public evaluations can be translated into targeted urban logistics strategies.

This study addresses that gap by examining citizen evaluations of a common cargo-bike delivery scenario across six European capital cities under a harmonized survey design. Its contribution lies not only in measuring acceptance, but in doing so in a way that makes evaluations comparable across cities and analytically decomposable into criterion-specific patterns. This makes it possible to distinguish between similar levels of overall acceptance and different underlying priority structures. In this sense, the study extends freight-perception research beyond identifying whether citizens respond positively or negatively to a logistics concept, toward understanding which service attributes structure those evaluations and how these structures vary across urban contexts.

For this purpose, we employ MUlticriteria Satisfaction Analysis (MUSA), a preference-disaggregation method designed to infer value systems from ordinal overall and criterion-level evaluations (Grigoroudis & Siskos, 2002, 2009). Originally developed in quality management to measure ex-post customer satisfaction, MUSA is used here

Table 1
Overview of related literature on urban freight participation, cargo-bike research, and public acceptance. The table highlights the gap addressed by this study: standardized, multi-city, multi-criteria citizen evaluation with decision-oriented outputs.

Stream	Focus	Methods	Scope	Main contribution and limitation
Participatory urban freight planning (Bjørgen et al., 2021; Lozzi et al., 2025; Maltese et al., 2023)	Governance and co-creation	Living labs, workshops, participatory processes	Mostly single-city	Offers governance and implementation insights, but usually lacks scalable and comparable measurement of citizen evaluations.
Citizen perceptions of urban freight (França et al., 2025; Parajuli & Pojani, 2018)	Freight externalities and support	Surveys, perception models	Single-city or limited comparison	Identifies salient freight concerns, but rarely evaluates a concrete service scenario or derives decision-oriented priorities.
Cargo-bike operations and impacts (Llorca & Moeckel, 2021; Rudolph & Gruber, 2017)	Feasibility, routing, emissions	Simulation, optimization, case studies, reviews	Pilot and operational settings	Shows where cargo bikes can work operationally, but not how citizens evaluate them across cities.
Cargo-bike enabling conditions and policy environment (Chatziioannou et al., 2024, 2026)	Safety, infrastructure, policy support	Reviews, expert-based assessment, policy analysis	Infrastructure and policy studies	Identifies conditions supporting cargo-bike deployment, but not large-scale citizen acceptance.
Common survey-analysis approaches in related literature (Amaya et al., 2021; Kiba-Janiak et al., 2024; Montero-Vega et al., 2025)	Ratings, effects, latent constructs, trade-offs	Descriptive statistics, ordered models, SEM, conjoint/discrete choice	Flexible, method-dependent	Useful for specific inferential tasks, but less suited to jointly deriving criterion importance and action-oriented diagnostics.
This study	Multi-city citizen evaluation of a cargo-bike delivery scenario	Standardized survey + MUSA	Six European capitals	Provides comparable acceptance levels, criterion weights, acceptance value functions, and decision-oriented diagnostics.

in a conceptually related but different way, namely to analyze ex-ante public acceptance of a hypothetical logistics service scenario. This adaptation is appropriate because, in both settings, respondents provide an overall judgment together with criterion-level evaluations, and the analytical objective is to infer how these partial evaluations combine into a global appraisal. In the present study, this global appraisal is interpreted as acceptance of the proposed service scenario rather than as realized satisfaction with an experienced service.

This framework is well suited to the survey design adopted here. Respondents assess the hypothetical service through ordered response categories rather than through repeated choice tasks or cardinal trade-off exercises. In contrast to descriptive summaries, MUSA relates criterion-level judgments to overall acceptance within a common framework. Compared with approaches frequently used in related studies, such as structural-equation models for citizen perceptions of urban freight (Amaya et al., 2021; França et al., 2025), ordered logit models for acceptance of last-mile delivery technologies (Montero-Vega et al., 2025), or conjoint designs for sustainable last-mile delivery preferences (Kiba-Janiak et al., 2024), MUSA yields outputs that are directly interpretable for policy design, including inferred criterion importance, marginal value functions, and action-oriented diagnostics. It is also attractive for large survey settings because the response task remains relatively simple, while still allowing the analyst to recover structured preference information from ordinal data. An additional advantage is that the method works directly with ordered evaluations and therefore does not impose equal distances between adjacent response categories. This is important in the present context, where moving from one response level to the next should not automatically be interpreted as a linear increase in perceived value (Grigoroudis & Siskos, 2009).

At the same time, the main analytical value of MUSA in this study is substantive rather than merely technical. The method makes it possible to identify not only which aspects of a cargo-bike delivery service are evaluated favorably, but also which of them are most influential in shaping overall acceptance and where improvement efforts are likely to be most effective. This provides a more policy-relevant reading of public opinion than overall averages alone. Table 1 summarizes the position of this study in relation to the most relevant strands of literature.

3. Methodology

To assess public acceptance of cargo bikes for last-mile delivery across multiple cities, we developed a structured yet accessible analytical framework. The basic idea is to present citizens with a concrete, evidence-based scenario of a cargo-bike delivery service and collect their evaluations of that service both overall and with respect to specific attributes. We then apply MUlticriteria Satisfaction Analysis (MUSA) to these ordinal evaluations to derive quantitative insights into public preferences. MUSA is an Multiple Criteria Decision Aiding (MCDA) method based on preference disaggregation that estimates a collective value structure from respondents' overall and criterion-level judgments. In the present study, this means identifying how strongly each criterion contributes to overall acceptance and how the service is evaluated on each criterion.

The method is particularly suitable here because the questionnaire elicits ordered evaluations rather than cardinal scores, and MUSA works directly with such ordinal information without assuming equal distances between adjacent response categories (Doumpos & Zopounidis, 2011; Grigoroudis & Siskos, 2002, 2009). At the same time, the model relies on an additive and compensatory representation of preferences, and the resulting estimates reflect stated evaluations of a hypothetical scenario rather than observed behavior. The outputs should therefore be interpreted as structured measures of scenario-based public acceptance under the assumptions of the model.

3.1. The MUSA formalism

MUSA is an ordinal regression method based on preference disaggregation. It infers a collective value system from respondents' global and criterion-level evaluations and represents this system through one global value function[1] and a set of partial value functions. Let M denote the number of respondents and let $G = \{g_1, \ldots, g_n\}$ denote the set of

[1] In the original MUSA formulation, the global and partial value functions represent satisfaction. In this study, the same mathematical structure is used, but these functions are interpreted as acceptance functions, since the empirical object of interest is public acceptance of a hypothetical cargo-bike delivery service.

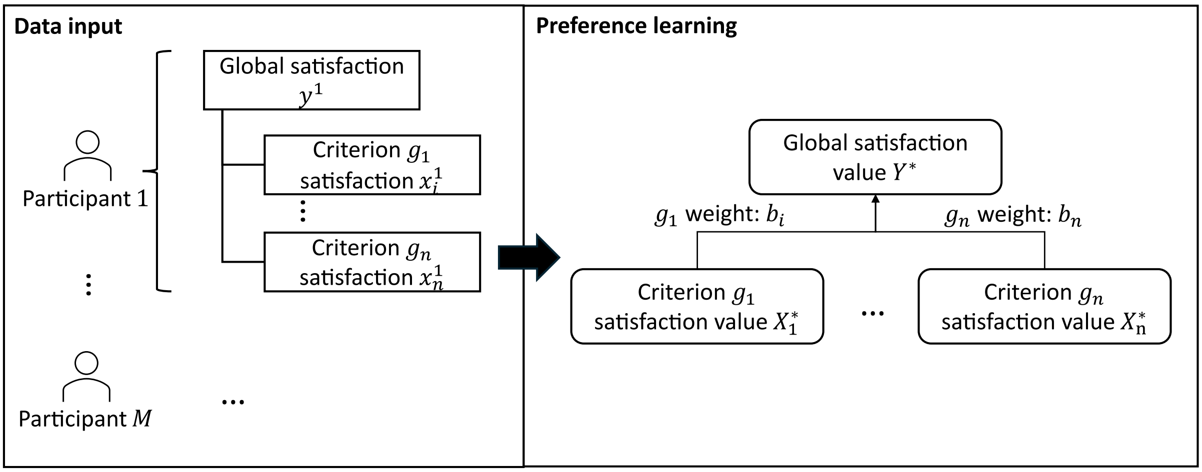


Fig. 1. Conceptual structure of the adapted MUSA model used in this study. Left: ordinal input data provided by respondents for overall acceptance and criterion-level acceptance evaluations of the hypothetical cargo-bike delivery service. Right: preference learning through a linear programming model that estimates monotone partial acceptance functions, a global acceptance function, and criterion weights. Following the original MUSA formalism, these value functions preserve the underlying mathematical structure of the method, but in the present study they are interpreted as acceptance functions rather than satisfaction functions. The outputs are normalized acceptance values and importance weights that support subsequent action and improvement analysis.

evaluation criteria, with $n = 5$ in this study: environmental impact, noise, traffic congestion, safety, and shipping cost.

Let

$$Y = \{y^1, \dots, y^\alpha\}$$

be the ordered set of global acceptance levels, and let

$$X_i = \{x_i^1, \dots, x_i^{\alpha_i}\}$$

be the ordered set of acceptance levels for criterion i. In our application, all scales are five-point ordinal scales, so that $\alpha = \alpha_i = 5$ for all i.

Fig. 1 summarizes the logic of the method. The model estimates a global acceptance value function Y^* and criterion-specific acceptance value functions X_i^*, all normalized on the interval $[0, 100]$. These functions map the ordinal response categories into latent value scores while preserving their ordering. Importantly, the distances between adjacent ordinal categories are not fixed a priori. Instead, they are inferred from the data through the estimated value functions. As a result, moving from one response level to the next does not need to correspond to a constant increase in latent acceptance. This is more consistent with the nature of ordinal scales, where category order is meaningful but equal spacing between categories cannot be assumed.

MUSA constructs a global value function Y^* and partial value functions X_i^* normalized within $[0, 100]$. The fundamental additive equation is:

$$\begin{cases} Y^* = \sum_{i=1}^{n} b_i X_i^* \\ \sum_{i=1}^{n} b_i = 1 \end{cases} \tag{1}$$

where b_i is the weight of criterion i. The value functions Y^* and X_i^* are piecewise linear, defined by the mapped scores of the ordinal levels. Specifically, $y^{*1} = x_i^{*1} = 0$ and $y^{*\alpha} = x_i^{*\alpha_i} = 100$.

To estimate these functions, MUSA employs a Linear Programming (LP) model. For each respondent $j \in \{1, \dots, M\}$, let y^{t_j} be their reported global acceptance and $x_i^{t_{ji}}$ be their reported acceptance for criterion i. The regression equation introduces double-error variables σ_j^+, σ_j^-:

$$Y_j^* = \sum_{i=1}^{n} b_i X_{ij}^* - \sigma_j^+ + \sigma_j^-, \tag{2}$$

where Y_j^* and X_{ij}^* are the values corresponding to the jth respondent's chosen levels. The objective of the LP is to minimize the total approximation error $F = \sum_{j=1}^{M} (\sigma_j^+ + \sigma_j^-)$. To linearize the problem, the model transforms the values into strictly positive steps between adjacent levels. We define:

$$\begin{aligned} z_m &= y^{*m+1} - y^{*m}, \quad \forall m = 1, \dots, \alpha - 1, \\ w_{ik} &= b_i x_i^{*k+1} - b_i x_i^{*k}, \quad \forall i,\ \forall k = 1, \dots, \alpha_i - 1. \end{aligned} \tag{3}$$

To ensure that higher declared acceptance corresponds to strictly higher value scores (monotonicity), we impose strictly positive thresholds γ, γ_i:

$$\begin{aligned} z_m \geqslant \gamma &\Leftrightarrow z'_m = z_m - \gamma \geqslant 0, \\ w_{ik} \geqslant \gamma_i &\Leftrightarrow w'_{ik} = w_{ik} - \gamma_i \geqslant 0. \end{aligned} \tag{4}$$

The final LP solves for z'_m, w'_{ik} and error terms σ to minimize F, subject to normalization constraints (summing to 100). the final LP model for MUSA is formulated as follows:

$$\begin{aligned} &\text{Minimize} \quad F = \sum_{j=1}^{M} (\sigma_j^+ + \sigma_j^-) \\ &\text{s.t.} \\ &\sum_{i=1}^{n} \sum_{k=1}^{t_{ji}-1} w'_{ik} - \sum_{m=1}^{t_j-1} z'_m - \sigma_j^+ + \sigma_j^- = \gamma \cdot (t_j - 1) - \sum_{i=1}^{n} \gamma_i \cdot (t_{ji} - 1), \\ &\quad \forall\ j = 1, 2, \dots, M, \\ &\sum_{m=1}^{\alpha-1} z'_m = 100 - \gamma \cdot (\alpha - 1), \\ &\sum_{i=1}^{n} \sum_{k=1}^{\alpha_i-1} w'_{ik} = 100 - \sum_{i=1}^{n} \gamma_i \cdot (\alpha_i - 1), \\ &z'_m, w'_{ik}, \sigma_j^+, \sigma_j^- \geqslant 0, \quad \forall m, k, j, i, \end{aligned} \tag{5}$$

where y^{t_j} and $x_i^{t_{ji}}$ are the judgments of the jth customer for global and partial acceptance with $y^{t_j} \in Y = \{y^1, \dots, y^\alpha\}$ and $x_i^{t_{ji}} \in X_i = \{x_i^1, \dots, x_i^{\alpha_i}\}$ for $i = 1, 2, \dots, n$.

3.2. Stability analysis and key indices

Because the MUSA linear program may admit multiple optimal or near-optimal solutions, we complement the basic estimation with a post-optimal stability analysis for evaluating robustness (Grigoroudis & Siskos, 2002; Voll et al., 2015). This is particularly relevant in large-scale survey settings, where several solutions may fit the observed evaluations almost equally well. Rather than relying on a single optimal solution, the robustness analysis explores the admissible solution space and identifies a representative set of criterion weights.

The procedure consists of two steps. First, the standard MUSA model in Eq. (5) is solved to obtain the minimum approximation error, denoted by F^*. A near-optimal feasible region is then defined by allowing solutions whose fit is only marginally worse:

$$F \leqslant F^* + \varepsilon, \tag{6}$$

where ε is a small tolerance parameter. Second, within this near-optimal region, additional LPs are solved to maximize and minimize the weight of each criterion b_i. These optimization runs identify the admissible intervals of the criterion weights without materially degrading model fit. The representative final solution is obtained by averaging the resulting admissible weights across these runs (Grigoroudis & Siskos, 2002).

The degree of variation across these solutions provides an indication of model stability. To summarize this variation, we use the Average Stability Index (ASI):

$$\mathrm{ASI} = 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{\sqrt{H\sum_{h=1}^{H}(b_i^h)^2 - \left(\sum_{h=1}^{H} b_i^h\right)^2}}{\frac{H}{n}\sqrt{n-1}}, \tag{7}$$

where H is the total number of post-optimal LPs, and b_i^h denotes the weight of criterion i obtained in the hth run. The ASI takes values in $[0, 1]$, with values close to 1 indicating high stability and lower values indicating greater dispersion of the admissible weights.

From the estimated value functions and representative weights, we derive four outputs that are central to the empirical analysis.

1. **Criterion weights** (b_i). These represent the relative importance of the evaluation criteria on a normalized scale such that $\sum_{i=1}^{n} b_i = 1$. They are recovered as:

$$b_i = \sum_{k=1}^{\alpha_i - 1} w_{ik}, \qquad \forall i \in \{1, \ldots, n\}. \tag{8}$$

2. **Global and partial acceptance indices** (S, S_i). These summarize the average acceptance level of the service overall and by criterion on a normalized 0–100 scale:

$$S = \frac{1}{100}\sum_{m=1}^{\alpha} p^m y^{*m}, \qquad S_i = \frac{1}{100}\sum_{k=1}^{\alpha_i} p_i^k x_i^{*k}, \tag{9}$$

where p^m denotes the observed proportion of respondents selecting global level m, and p_i^k denotes the proportion selecting level k for criterion i. Higher values indicate more favorable evaluations.

3. **Demanding indices** (D, D_i). These indices, defined on $[-1, 1]$, characterize the curvature of the estimated value functions. Positive values indicate a relatively demanding evaluation profile, in the sense that higher verbal ratings correspond to disproportionately larger increases in latent acceptance. Negative values indicate the opposite.

4. **Improvement indices** (I_i). These combine criterion importance and current performance:

$$I_i = b_i(1 - S_i), \qquad \forall i \in \{1, \ldots, n\}. \tag{10}$$

A high value of I_i identifies a criterion that is both important and currently underperforming, and therefore represents a promising leverage point for policy intervention or communication.

3.3. Application framework

Applying MUSA in a mass-participation context (our survey has around 2500 respondents) and across multiple cities presents challenges. The original MUSA method has mostly been used on smaller samples. With a very large and heterogeneous sample, directly performing MUSA could obscure distinct preference patterns if sub-groups differ greatly. Moreover, comparing MUSA results between cities is not straightforward if each city's respondents evaluated a different service context. We addressed these issues through two key adaptations: (1) segmenting the data by relevant clusters to ensure more internally consistent groups for analysis, and (2) using a standardized hypothetical service scenario for all cities to provide a common evaluation baseline.

The overall framework for our methodology can be summarized in the following steps (these correspond to the flow depicted in Fig. 2):

1. **Scenario Development and Questionnaire Design:** We developed a detailed hypothetical cargo bike delivery service scenario, including key performance criteria (environmental impact, noise level, etc.). Based on this scenario, we designed a questionnaire that first explains the service and its criteria, then asks respondents to rate their acceptability with each criterion and overall. Additional questions capture socio-demographic data and personal context (e.g., online shopping habits, familiarity with cargo bikes).
2. **Survey Deployment and Data Collection:** The questionnaire was administered online to residents in the target cities. To maximize reach and diversity, the survey was made available in local languages (English, French, Italian, Polish, Dutch where appropriate) and distributed through a panel provider. We collected responses from London, Paris, Rome, Dublin, Warsaw, and Brussels.
3. **Data Preprocessing and Clustering:** We cleaned and formatted the data, ensuring that all acceptance responses were on a unified 5-point Likert scale. We then examined the data for significant differences across demographic groups within each city. A simple statistical test was used on overall acceptance and criterion scores to check, for example, if age or area of residence led to different mean responses. Where a clear significant split was found (e.g., urban vs. suburban in London), we segmented the dataset accordingly. This step ensures that the subsequent MUSA analysis can be done on relatively homogeneous sub-samples, aligning with MUSA's methodological assumptions.
4. **MUSA Application for Each Group:** We applied the MUSA model separately to each cluster (or city, if no clustering was needed within a city). In each case, respondents' overall acceptance with the service is the dependent variable, and their acceptance ratings on the individual criteria are the inputs for the ordinal regression-based preference disaggregation.
5. **Strategy Derivation and Comparison:** From the MUSA results of each cluster, we identified which criteria are strengths (high acceptance and importance) and which are weaknesses or improvement areas (high importance but low acceptance). We then compared results across the cities/clusters to see how priorities differ. Finally, we synthesized overarching insights that apply broadly and could inform multi-city strategies.

4. Implementation: Survey design and data collection

To operationalize the methodology, we administered a mass-participation online survey centered on a hypothetical cargo-bike delivery service, *EcoDeliver*. The scenario was designed to be realistic while remaining sufficiently standardized for use across all study cities. Its key performance attributes were defined based on recent literature and current industry practice. Table 2 summarizes the five criteria presented to respondents — environmental impact (CO_2), noise, traffic/congestion, safety, and shipping cost — together with concise evidence-based descriptions. These criteria reflect the main perceived benefits and trade-offs emphasized in the acceptance literature, including environmental improvements, social impacts, and individual

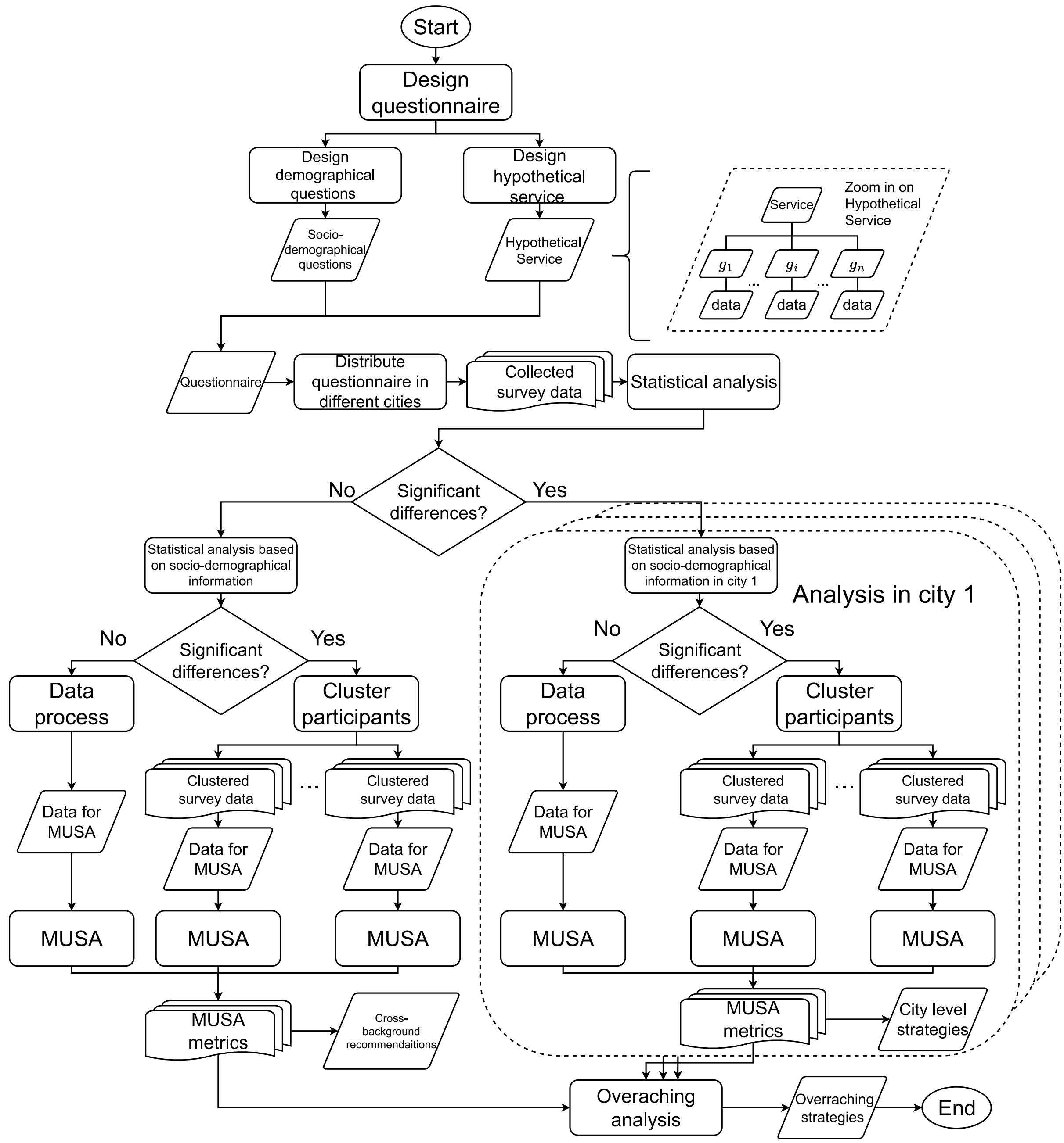


Fig. 2. Multi-city acceptance elicitation framework for cargo bike delivery acceptance (survey design, clustering, MUSA analysis, strategy development).

cost considerations. Each criterion vignette included a quantitative or factual anchor, such as “around 70%–90% lower CO_2 than diesel vans”, to ground respondents’ judgments in concrete terms and reduce information asymmetry. Immediately after each vignette, respondents were asked to evaluate the scenario on a five-point Likert scale: *Extremely unacceptable*, *Somewhat unacceptable*, *Neither acceptable nor unacceptable*, *Somewhat acceptable*, and *Extremely acceptable*. A final overall item then asked: “*Considering all of the above, how acceptable is EcoDeliver as a last-mile delivery option in your city?*” All MUSA inputs are based on this five-point ordinal acceptance scale, which allows the estimated value functions and weights to be interpreted as ex-ante public acceptance.

In addition to the acceptance items, the survey included a set of sociodemographic and background questions (see Table 3). These variables serve both to characterize the sample and to support subgroup analyses. Respondents were asked about their familiarity with the concept of cargo-bike delivery, frequency of online shopping, age, gender, education level, employment status, and whether they lived in an urban, suburban, or rural area. We also included a question on respondents’ satisfaction of their current living conditions, measured on the same five-point scale, as a general indicator of how they perceived their local urban environment. This variable provides contextual information that may help explain variation in acceptance of the proposed service scenario. The survey instrument was reviewed by experts and piloted with a small group to ensure that the wording was clear and that the scenario was understandable. For completeness, the full questionnaire is provided in the supplementary material.

We deliberately focused on six European capitals: London, Paris, Rome, Dublin, Warsaw, and Brussels, to balance comparability (all are national capitals with broadly similar last-mile needs and regulatory levers) and contrast (they differ sharply in policy, urban form, and logistics baselines). First, the set spans Western, Southern, and Central–Eastern Europe, ensuring regional diversity in travel cultures, income

Table 2
Key criteria and scenario information of hypothetical cargo bike delivery service *EcoDeliver* presented to respondents (excerpt from survey). Each criterion was followed by a acceptance question on a 5-point scale.

Code	Criterion	Description (Scenario information)
g_1	CO_2 emissions	Cargo bike delivery can significantly reduce emissions. For example, *EcoDeliver* is projected to cut carbon emissions by 70%–90% per package compared to diesel vans (and about one-third compared to electric vans). This means cleaner air and lower greenhouse gases in your city.
g_2	Noise	*EcoDeliver* bikes operate without engine or exhaust noise and are markedly quieter in pass-by than motor vehicles. Current EU limits for new vehicles allow *around* 70 dB(A) for vans and 77–80 dB(A) for mopeds/motorcycles; e-bikes typically produce only tyre/chain and wind noise, often near ambient urban levels. This can translate into quieter streets during deliveries.
g_3	Traffic	Using cargo bikes can reduce traffic congestion. They take up far less road space than vans; our *EcoDeliver* bikes need roughly 75% less road space per delivery than a car. Couriers can also use bike lanes and avoid adding to car traffic. This improvement means fewer delivery vans clogging city streets, potentially leading to smoother traffic flow.
g_4	Safety	Cargo bikes pose lower risk to pedestrians and cyclists than large vehicles, and they travel at lower speeds (max 25 km/h). *EcoDeliver* is designed with safety in mind — fewer heavy vans blocking roads and sidewalks, and trained riders following safe practices. This could reduce the number and severity of accidents related to deliveries in the city.
g_5	Shipping cost	Because *EcoDeliver* invests in fair wages and good working conditions for bike couriers, the delivery fee might be slightly higher. We estimate a 10%–20% increase in shipping cost compared to standard delivery services. In return, you support sustainable jobs and cleaner deliveries.

Table 3
Socio-demographic and background questions (SDQ) included in the survey.

Question (variable code)	Response options
How would you describe your understanding of the term "cargo bike delivery"? (*SDQ1: Understanding*)	Fully knowledgeable/Somewhat knowledgeable/Not very knowledgeable/Not at all knowledgeable
How often do you shop online for clothing and fashion products? (*SDQ2: Online shopping frequency*)	Daily/Weekly/Monthly/Rarely/Never
What is your age? (*SDQ3: Age*)	18–24/25–34/35–44/45–54/55 or above
What is your gender? (*SDQ4: Gender*)	Male/Female
What is your highest level of education? (*SDQ5: Education*)	High school or less/College or vocational/Bachelor's degree/Postgraduate (Master's/PhD)
Are you currently employed? (*SDQ6: Employment*)	Yes, full-time/Yes, part-time/No, not employed/Student/Retired
What type of area do you live in? (*SDQ7: Area*)	Urban (city center)/Suburban (outskirts of city)
How satisfied are you with your current living conditions (neighborhood, city)? (*SDQ8: Life satisfaction*)	5-point scale: Extremely satisfied to Extremely dissatisfied

Note: All socio-demographic questions have an option "Prefer not to say".

levels, and street typologies, so that acceptance drivers can be stress-tested across very different built environments. Second, these cities sit at different points on the policy and infrastructure spectrum relevant to cargo bikes (e.g., low-emission and access-managed zones, cycling networks, curb and loading management, consolidation hubs), which creates meaningful variation in what residents value (environment, noise, congestion, safety, cost) and how strongly those criteria shape overall acceptance. Third, they represent a wide range of logistics market conditions — high-volume e-commerce corridors, tourism-heavy centers, and government/administrative districts — making findings informative for both public and private actors. Practical considerations also mattered. As capitals, these cities offer accessible sampling frames and comparable governance instruments needed to translate acceptance into implementation.

We took care to translate the survey into local languages where needed (French in Paris, Italian in Rome, Polish in Warsaw, both French and Dutch for Brussels) to reduce language barriers. All versions were kept as equivalent as possible in wording and content.

4.1. Statistical analysis

Data were collected through a professional online panel provider in two fieldwork waves conducted in 2024 and 2025. The questionnaire was implemented in Qualtrics. Because recruitment difficulty varied across the six countries, the survey followed a quota-guided nonprobability design with soft city targets rather than equal allocation. The initial targets were approximately 700 respondents for London, 550 for Paris, 500 for Rome, 250 for Dublin, and 200 each for Brussels and Warsaw. Because Brussels and Warsaw did not reach their targets in the first wave, a second wave was conducted for these two cities. Within each city, soft balancing quotas were applied for gender and broad age groups, while area of residence was monitored but not subject to a hard quota. Minor deviations from these targets were allowed when necessary to complete fieldwork in smaller or harder-to-recruit panels.

The questionnaire was administered in local-language versions while retaining the same core structure across cities: informed consent, a city screening item, the standardized *EcoDeliver* scenario, five criterion-level acceptance items, one overall acceptance item, and background questions. The full questionnaire is provided in the supplementary material. The panel provider reported the number of eligible survey starts, completed questionnaires, and quality-control exclusions. Fieldwork produced 3222 eligible starts and 2641 completed questionnaires. After excluding 126 responses during data-quality screening, the final analytic sample comprised 2515 respondents. This corresponds to an overall completion rate of 82.0% among eligible starts and a retention rate of 95.2% from completed questionnaires to the final analytic sample. Exclusions were based on duplicate identifiers or repeated device/IP matches, as well as failed attention or straightlining checks. Average completion times ranged from 463 s in Warsaw to 675 s in Paris. No post-stratification weights were applied, since the purpose of the study is to compare acceptance structures across city samples. The resulting dataset should therefore be interpreted as a quota-guided online panel sample rather than a statistically representative sample of each city population.

The final pooled analytic sample consists of $N = 2515$ respondents across six capitals: London ($n = 724$, 28.8%), Paris ($n = 571$, 22.7%), Rome ($n = 535$, 21.3%), Dublin ($n = 278$, 11.1%), Brussels ($n = 204$, 8.1%), and Warsaw ($n = 203$, 8.1%). Across cities, the share of female respondents ranged from 47.5% in Brussels to 63.7% in Dublin, while the proportion living in urban cores ranged from 58.3% in Dublin to 83.3% in Warsaw. The most common age groups were 25–34 and 35–44, with Paris showing a somewhat younger profile. A detailed description of the sample is reported in Table 4.

Acceptance ratings were collected on a five-point ordinal scale for each criterion (CO_2, noise, traffic/congestion, safety, and shipping cost) and for the overall evaluation item. Because these responses are ordinal and may deviate from normality, all subsequent analyses rely on nonparametric procedures. We first report descriptive statistics by country, including medians and interquartile ranges, together with a compact profile of the city samples (Tables 4 and A.7).

To test for cross-country differences, we applied the Kruskal–Wallis (KW) test to each acceptance variable, reporting the H statistic, degrees of freedom, p-value, and the KW effect size $\hat{\varepsilon}^2$ (Table A.8) (Smeeton et al., 2025). For variables with a significant KW result ($\alpha = 0.05$), we conducted pairwise Mann–Whitney rank-sum tests with Holm correction for multiple comparisons and report significant country pairs with adjusted p and Cliff's δ as an effect size (Table A.9) (Mann & Whitney, 1947).

We then examined socio-demographic influences on overall acceptance in the pooled sample using KW tests across groups for gender, age, education, employment status, residential area (urban/suburban), cargo-bike knowledge, online shopping frequency, and current living condition (Table A.10).

The key findings are as follows:

1. *Cross-country:* Traffic/congestion ($p < 0.001$), safety ($p < 0.001$), CO_2 ($p = 0.023$), and noise ($p = 0.049$) showed significant between-country differences (small effects: $\hat{\varepsilon}^2 \in [0.002, 0.012]$), whereas overall acceptance and shipping cost did not differ significantly (Table A.8). Post-hoc tests indicated, for example, London > Paris for traffic (Holm-adjusted $p < 0.001$, Cliff's $\delta = 0.153$) and Paris < Rome for traffic ($p < 0.001$, $\delta = -0.158$); London > Paris for safety ($p < 0.001$, $\delta = 0.139$); and Paris < Rome for CO_2 ($p = 0.017$, $\delta = -0.102$) (Table A.9). Medians and IQRs by country are reported in Table A.7.
2. *Socio-demographics (pooled):* Overall acceptance increased monotonically with self-reported living condition (KW $p < 10^{-23}$, $\hat{\varepsilon}^2 = 0.044$) and cargo-bike knowledge (KW $p < 10^{-21}$, $\hat{\varepsilon}^2 = 0.039$); online shopping frequency and residential area also showed smaller but significant effects (Table A.10). In contrast, age, gender, and education did not exhibit significant differences in overall acceptance.

Table 4
Sample size and selected socio-demographic features by country.

Country	N	Female (%)	Urban (%)	Most common age
Brussels	204	47.5	72.1	35–44
Dublin	278	63.7	58.3	35–44
London	724	53.2	73.1	25–34
Paris	571	61.1	79.0	18–24
Rome	535	60.9	79.3	35–44
Warsaw	203	57.6	83.3	25–34

Differences between countries are statistically detectable but small in magnitude, suggesting broadly similar acceptance with city-specific emphases (stronger perceived gains in traffic and safety in some capitals). Knowledge about cargo-bike delivery and satisfaction with one's living context are the strongest correlates of higher overall acceptance, highlighting the value of awareness campaigns and visible, local quality-of-life improvements.

Following the acceptance elicitation framework, we first assessed whether any capitals could be pooled into a single "city group" with a common model. Although overall acceptance levels were broadly comparable across cities, the underlying drivers, i.e., criteria, differed enough that pooling would obscure city-specific priorities; we therefore retained a separate model for each city.

We then tested, within each city, whether respondents living in Urban (city-center) and Suburban (outskirts) areas should be modeled separately. For each city, we compared overall acceptance using a two-sided Mann–Whitney test, applied Holm correction across the six city-wise comparisons, and recommended a split only if three conditions were met simultaneously: $p_{\text{Holm}} < 0.05$, $|\delta| \geq 0.10$, and both subgroups had $n \geq 50$. The resulting decisions are:

- **London:** *Split.* This is the only city that meets all three criteria ($n_{\text{urban}} = 529$, $n_{\text{suburban}} = 195$, $p_{\text{Holm}} = 0.0037$, Cliff's $\delta = 0.152$). We therefore estimate two MUSA models (London–Urban and London–Suburban).
- **Paris:** *Do not split.* Although subgroup sizes are adequate ($n_{\text{urban}} = 451$, $n_{\text{suburban}} = 120$), the estimated difference is below the practical threshold and not significant after correction ($p_{\text{Holm}} = 0.716$, $\delta = 0.065$).
- **Rome:** *Do not split.* Subgroup sizes are also adequate ($n_{\text{urban}} = 424$, $n_{\text{suburban}} = 111$), but the Urban–Suburban difference is essentially null ($p_{\text{Holm}} = 1.000$, $\delta = -0.009$).
- **Dublin:** *Do not split.* Both subgroups are large enough ($n_{\text{urban}} = 162$, $n_{\text{suburban}} = 116$), yet the difference is negligible ($p_{\text{Holm}} = 1.000$, $\delta = 0.034$).
- **Warsaw:** *Do not split.* The estimated gap is not significant after correction ($p_{\text{Holm}} = 0.359$, $\delta = 0.179$) and the suburban subgroup is too small for stable two-model estimation ($n_{\text{suburban}} = 34 < 50$; $n_{\text{urban}} = 169$).
- **Brussels:** *Do not split.* The suburban subgroup is just above the minimum size ($n_{\text{urban}} = 147$, $n_{\text{suburban}} = 57$), but the adjusted test is not significant ($p_{\text{Holm}} = 0.716$), so the evidence for a separate model remains insufficient despite a small effect estimate ($\delta = 0.112$).

5. Results

Following the grouping procedure reported in the previous section, MUSA was applied to seven groups: London–Urban, London–Suburban, Paris, Rome, Dublin, Warsaw, and Brussels. The same implementation settings were used for all groups to ensure comparability. The models were run in Python 3.9.13 using `gurobipy` with Gurobi 10.0.1. All global and criterion-level evaluations were coded on the same five-point ordinal scale. Strict monotonicity of the value functions was

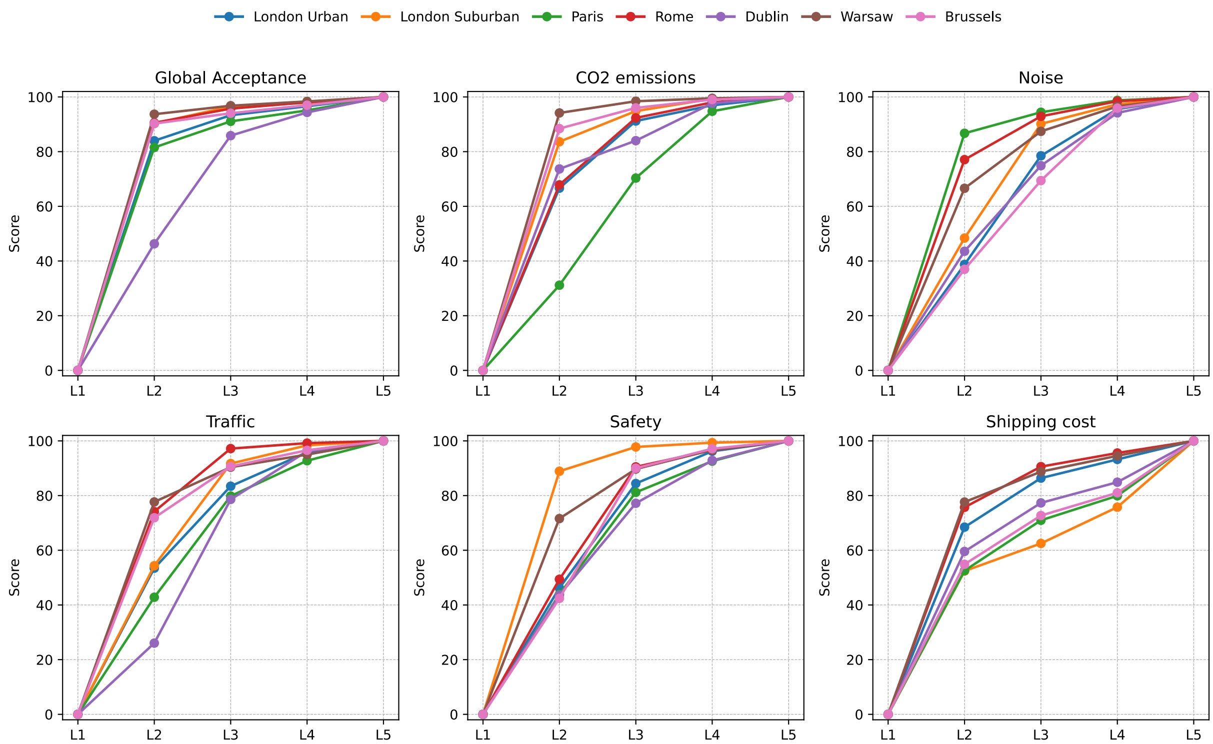


Fig. 3. Learned acceptance functions for each city/group. Scores are normalized on a 0–100 scale and increase monotonically across the five ordered response levels.

enforced using $\gamma = \gamma_i = 0.01$. The post-optimal analysis was conducted in a near-optimal region defined by

$$F \leqslant F^* + \varepsilon, \qquad \varepsilon = 0.1F^*,$$

and the reported post-optimal values were obtained by averaging the solutions from the corresponding near-optimal extreme-point runs.

The robustness analysis indicates that the resulting preference structures are stable across all seven groups. The ASI values are above 0.85 in every case and exceed 0.90 for Brussels and Warsaw. This indicates that the admissible near-optimal solutions remain tightly concentrated, so the reported weights can be interpreted as robust summaries of each group's acceptance structure rather than as artifacts of a single optimal solution.

Fig. 3 presents the learned global and criterion-level acceptance functions for all cities/groups. Each function is normalized on a 0–100 scale and is strictly increasing across the five ordered response levels. Their shapes show how respondents translate the ordinal response categories into latent acceptance values. In particular, the functions confirm that adjacent verbal categories are not treated as equally spaced: some transitions correspond to larger increases in latent acceptance than others. These learned functions provide the basis for the acceptance indices, criterion weights, demanding indices, and improvement indices. These metrics are recorded in Table 5, which reports the main MUSA outputs by city/group: overall acceptance S and demanding index D, criterion-level acceptance S_i, criterion weights b_i, demanding indices D_i, and improvement indices I_I. Together, these results show not only how positively the hypothetical *EcoDeliver* scenario was evaluated, but also which criteria structure that evaluation, how respondents use the ordinal response scale, and where the greatest room for improvement remains.

Overall acceptance is high in all seven groups, ranging from 93.1 in Dublin to 98.3 in Warsaw. Criterion-level acceptance is likewise high for most attributes. Environmental impact records consistently strong acceptance across all groups, with values from 93.5 in Paris to 99.6 in Warsaw. Noise and traffic are also very positively evaluated throughout, and safety remains above 89 in all groups. By contrast, shipping cost is the least favorably evaluated criterion in every group and shows the widest dispersion, ranging from 67.1 in London (Suburban) to 90.4 in Rome. Thus, the scenario is broadly well received across cities, but the cost component remains the most fragile element of the evaluation.

The importance weights reveal clearer cross-city differences than the acceptance scores themselves. Environmental impact receives the highest weight in London (Urban), Dublin, Warsaw, and Brussels, and is jointly dominant with safety in London (Suburban). Paris stands out because noise is by far the most influential criterion ($b = 0.531$), while Rome is characterized by the strongest emphasis on traffic/congestion ($b = 0.395$). These contrasts indicate that broadly similar overall acceptance levels may rest on different local priority structures. In other words, citizens in different cities may all evaluate the same scenario favorably, but not for the same reasons. The demanding indices add an important qualification to these favorable evaluations. All global demanding indices are negative, from −0.510 in Dublin to −0.926 in Warsaw, indicating relatively lenient evaluation profiles under the presented scenario. At the criterion level, the same pattern largely holds: all demanding indices remain negative, but their magnitude varies by attribute and city. For instance, cost is evaluated comparatively less leniently in London (Suburban) ($D = -0.271$) and Paris ($D = -0.356$) than in Rome ($D = -0.746$), while environmental impact is evaluated very leniently in Warsaw ($D = -0.947$) and Brussels ($D = -0.890$). These results suggest that the high acceptance levels partly reflect the favorable framing of the scenario. Nevertheless, it should be interpreted as acceptance of the specific *EcoDeliver* proposition rather than unconditional support for cargo-bike delivery in general.

The improvement indices identify where additional gains in acceptance remain most plausible within the stated scenario. Shipping cost emerges as the most consistent improvement priority: it has the highest I_i in Paris (0.021), Rome (0.006), Dublin (0.026), Brussels (0.013), and London (Suburban) (0.012), and is tied for the highest value in London (Urban) (0.013) and Warsaw (0.005). This pattern is important because it combines relatively weaker acceptance with sufficient importance to matter for overall evaluations. By contrast, environmental impact often carries the greatest weight but has less

Table 5
MUSA metrics by city/group: overall acceptance (S), criterion-level acceptance (S_i), importance weights (b_i), demanding indices (D_i), and improvement indices (I_i).

City/Group	Metric	Overall	Criterion				
			CO_2	Noise	Traffic	Safety	Cost
London (Urban)	S	95.9	97.4	95.4	96.0	95.5	86.6
	b	–	0.490	0.120	0.138	0.156	0.096
	D	−0.826	−0.699	−0.419	−0.551	−0.512	−0.654
	I	–	0.013	0.006	0.006	0.007	0.013
London (Suburban)	S	97.4	99.1	97.2	97.3	99.3	67.1
	b	–	0.358	0.097	0.151	0.358	0.037
	D	−0.898	−0.852	−0.572	−0.630	−0.907	−0.271
	I	–	0.003	0.003	0.004	0.003	0.012
Paris	S	94.3	93.5	98.0	91.0	90.9	77.0
	b	–	0.159	0.531	0.110	0.109	0.091
	D	−0.785	−0.308	−0.866	−0.436	−0.451	−0.356
	I	–	0.010	0.011	0.010	0.010	0.021
Rome	S	97.9	98.4	98.5	99.1	96.3	90.4
	b	–	0.167	0.272	0.395	0.100	0.065
	D	−0.895	−0.721	−0.790	−0.803	−0.577	−0.746
	I	–	0.003	0.004	0.004	0.004	0.006
Dublin	S	93.1	97.3	93.6	94.1	89.8	80.0
	b	–	0.387	0.128	0.230	0.123	0.132
	D	−0.510	−0.703	−0.417	−0.339	−0.424	−0.479
	I	–	0.011	0.008	0.013	0.013	0.026
Warsaw	S	98.3	99.6	95.4	95.6	94.7	89.7
	b	–	0.701	0.072	0.114	0.067	0.046
	D	−0.926	−0.947	−0.670	−0.754	−0.720	−0.740
	I	–	0.003	0.003	0.005	0.004	0.005
Brussels	S	96.6	99.1	95.6	95.4	95.0	76.7
	b	–	0.529	0.125	0.144	0.145	0.058
	D	−0.875	−0.890	−0.349	−0.727	−0.530	−0.390
	I	–	0.005	0.005	0.007	0.007	0.013

Note: S is reported on a 0–100 scale; b, D, and I are the normalized MUSA indices used in the action and improvement diagrams.

residual room for improvement because its acceptance is already very high. Hence, the results distinguish between criteria that primarily confer legitimacy on the scenario, such as environmental impact in several groups, and criteria that offer the clearest leverage for further gains, most notably shipping cost.

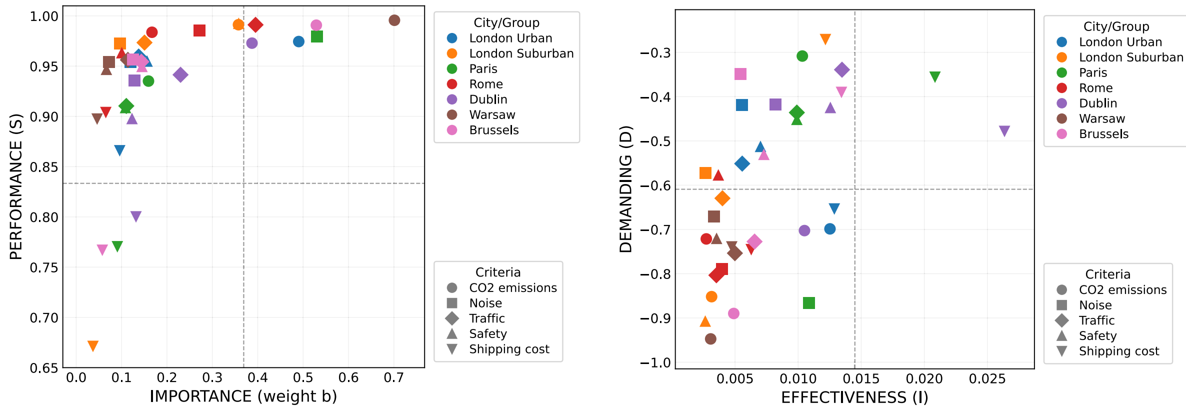
Taken together, the results point to a common structure with local variation. Across all groups, the environmental, noise, traffic, and safety dimensions are generally viewed positively, while shipping cost is the most contested attribute. However, the criterion weights, demanding indices, and improvement indices show that the relative role of these attributes differs across cities. The substantive implication is that similar overall acceptance levels should not be read as evidence of homogeneous public priorities.

To synthesize these results visually, Fig. 4 presents the two standard diagnostic views derived from MUSA: the action diagram and the improvement diagram. The action diagram in Fig. 4(a) combines the criterion weights (b_i) and acceptance indices (S_i), and therefore identifies which criteria are both influential in shaping overall acceptance and comparatively weaker in current performance. The improvement diagram in Fig. 4(b) complements this perspective by plotting the improvement indices (I_i) against the demanding indices (D_i), thereby showing where additional gains in acceptance are most likely and whether the observed evaluations arise under relatively lenient or more exacting response profiles. Read together, the two diagrams distinguish between established strengths of the scenario and the criteria that remain the most relevant priorities for improvement.

A first overall observation from the action diagram is that most criteria are located in the upper part of the plot, which confirms that the hypothetical *EcoDeliver* scenario was evaluated favorably across cities/groups. The diagram therefore does not reveal a classic pattern of severe underperformance on highly important criteria. Instead, it highlights different city-specific anchors of acceptance. In London (Urban), environmental impact clearly dominates, with traffic and safety occupying secondary positions. In London (Suburban), environmental impact and safety jointly structure the evaluation. Paris is distinguished by the primacy of noise, while Rome is characterized by the particularly strong role of traffic/congestion, followed by noise. Dublin shows a two-pillar structure centered on environmental impact and traffic, whereas Warsaw and Brussels are strongly anchored in environmental performance. In this sense, the action diagram indicates that the main "action" in most groups is less about correcting a major weakness than about preserving and reinforcing the criterion that already carries the local acceptance structure. The improvement diagram refines this reading by showing which criteria still offer the greatest room for additional gains. Shipping cost is the most consistent improvement priority across groups. It is the rightmost or among the rightmost criteria in Paris, Dublin, Brussels, and both London groups, indicating that it offers the strongest residual leverage for increasing acceptance. This is especially pronounced in Dublin and Paris, where cost combines comparatively modest acceptance with sufficient importance to matter for the overall evaluation. In Brussels and suburban London, cost is also the clearest remaining improvement priority, even though it is not the dominant driver of acceptance. By contrast, in Rome and Warsaw all improvement indices are low, which suggests that within the stated scenario there is relatively limited additional room to raise acceptance through marginal improvements to individual criteria.

6. Discussion

The results suggest that policy implementation should not rely on a single narrative of cargo-bike adoption. Although overall acceptance of the hypothetical *EcoDeliver* scenario is high across all groups, the MUSA weights show that the criteria structuring that acceptance differ by city. The practical implication is therefore not that each city requires a wholly different policy regime, but that existing SUMPs and SULPs



(a) Action diagram: importance vs. acceptance. (b) Improvement diagram: effectiveness vs. demandingness.

Fig. 4. Diagnostic diagrams for all cities/groups derived from the MUSA analysis.

instruments should place different emphasis on different dimensions of the same transition. Within the scope of the present study, the findings are best interpreted as guidance on which local co-benefits are likely to resonate most strongly with citizens when cargo-bike delivery is introduced or expanded.

- **London.** The split between the urban and suburban groups suggests that a uniform metropolitan message would miss an important within-city distinction. In the urban core, acceptance is primarily anchored in environmental impact, whereas in the suburban group safety becomes equally influential. This pattern fits well with London's current policy landscape, in which freight decarbonization and cargo-bike deployment are already embedded in the Freight and Servicing Action Plan and the Cargo Bike Action Plan (Transport for London, 2023; Transport for London, 2019). At the same time, the broader transport context of low-traffic and low-emission measures has also sharpened attention to street design and road safety (Mayor of London, 2018; Yang et al., 2022). Accordingly, in central areas cargo-bike delivery may be communicated mainly through emissions reduction and cleaner freight, whereas in outer areas implementation is more likely to benefit from visible safety measures, including protected cycling links, safe curbside operations, and clear operating rules for interactions with mixed traffic.
- **Paris.** Paris stands out because noise is the dominant criterion in the acceptance structure, while environmental impact, traffic, and safety all remain positively evaluated but secondary. This suggests that, in the Paris context, the public value of cargo-bike delivery may be strengthened less by decarbonization messaging alone than by emphasizing quieter operations and reduced disturbance in dense mixed-use neighborhoods. This reading is consistent with the wider policy setting in Paris, where recent cycling policies have altered street use and vehicular flow (Natterer et al., 2026), and where the city's cycling plan explicitly links cycling infrastructure, urban quality, and the transition toward cleaner mobility (Ville de Paris, 2021). Within such a framework, implementation choices such as low-noise delivery practices, careful management of loading and unloading activity, and continued integration of cargo bikes into broader cycling and climate strategies appear especially relevant.
- **Rome.** Rome presents a different pattern, with traffic/congestion clearly the most influential criterion. In this case, the most relevant implementation narrative is likely to be the spatial and operational efficiency of cargo bikes rather than environmental benefit alone. This interpretation is coherent with Rome's PUMS-related urban logistics actions, which place strong emphasis on logistics management, access regulation, and zero-emission freight strategies (Roma Servizi per la Mobilità, 2024). The results therefore suggest that communication and policy design in Rome should stress the visible reduction of delivery traffic in constrained urban areas, especially where freight competes with limited street space. Measures linked to urban consolidation, restricted-traffic areas, and the substitution of van-based last-mile trips appear particularly well aligned with the local acceptance structure.
- **Dublin.** Dublin shows a more balanced structure, but environmental impact and traffic are the two leading criteria, while shipping cost has the highest improvement index. This suggests that implementation in Dublin may benefit from a dual emphasis: first, making the environmental and traffic benefits of cargo-bike delivery visible; second, limiting or clearly justifying any cost premium. This interpretation aligns with the Dublin City Centre Transport Plan, whose technical work on goods movement explicitly discusses distribution centers, freight consolidation, and the role of cargo bikes in reducing pressure from heavy urban freight traffic (Dublin City Council and National Transport Authority, 2024). In practical terms, the results point toward infrastructure and logistics arrangements that can reduce delivery costs while keeping the environmental and congestion benefits salient.
- **Warsaw and Brussels.** Warsaw and Brussels are the two groups in which environmental impact is most clearly dominant, although the improvement analysis shows that cost remains the main residual weakness in Brussels and one of only limited remaining weaknesses in Warsaw. In Warsaw, this strong environmental emphasis is consistent with the broader policy salience of low-emission measures and sustainable mobility planning (Holnicki et al., 2023; Warsaw Functional Area, 2023). In Brussels, it also fits a context in which air-quality and low-emission policies have become increasingly prominent (Brussels Capital Region, 2021; Bruyneel et al., 2025). A cautious reading is therefore that environmental performance is the most salient source of legitimacy for cargo-bike delivery in these two contexts, while pricing remains a practical consideration rather than the main organizing principle of acceptance. For policy implementation, this suggests linking cargo-bike initiatives closely to clean-air, climate, and sustainable-mobility objectives, while ensuring that operational support prevents cost barriers from becoming more prominent than they currently appear.

Across all groups, one recurrent result deserves particular attention: shipping cost is the most consistent improvement priority. This does not imply that citizens are generally insensitive to price, nor does it justify the stronger claim that cost is unimportant. Rather, within the presented scenario, cost is typically the criterion where evaluations are somewhat weaker while still retaining enough importance to matter. For practitioners, this points toward a practical recommendation: where cargo-bike delivery is introduced, implementation should aim either to keep any price premium modest or to make the associated value proposition especially clear. Operational measures that improve efficiency, such as micro-hubs, better curbside organization, or integration with existing low-emission freight strategies, are therefore relevant not only for logistics performance but also for public acceptance.

More broadly, the results support three cross-cutting recommendations for city managers and logistics stakeholders. First, implementation should be linked to the locally dominant public priority identified by the city-specific model, rather than communicated through a uniform cross-city narrative. Second, cost should be monitored as the most likely source of residual resistance, even where it is not the dominant driver of acceptance. Third, cargo-bike policies should be embedded within existing SUMPs and SULPs instruments rather than treated as standalone pilot measures, because public acceptance is likely to depend not only on the vehicles themselves but also on how well they fit within the wider urban mobility environment.

6.1. Limits of interpretation and sensitivity to scenario framing

These implementation implications should, however, be read within the limits of the study design. They are derived from a hypothetical scenario and therefore identify differentiated priorities in public acceptance structures rather than definitive causal effects of existing policies. In particular, the *EcoDeliver* scenario was framed in favorable terms across several dimensions, including substantial environmental gains, quieter operation, reduced congestion, improved safety, and only a moderate increase in shipping cost. Under such conditions, high acceptance scores are not unexpected.

For this reason, the results should not be interpreted as evidence of broad or unconditional public acceptance of cargo-bike delivery in general. Rather, the main contribution of the study lies in showing how citizens in different cities prioritize the attributes of a favorable hypothetical cargo-bike delivery service. In practical terms, the policy relevance of the study lies less in the absolute level of acceptance and more in the relative importance of environmental impact, noise, traffic, safety, and cost. These priority structures remain informative for urban planning because they indicate which dimensions are most likely to shape public evaluations if a cargo-bike service is introduced or expanded in a given context.

To probe how sensitive these priority structures are to the favorable framing of the scenario, we conducted an additional small sensitivity survey in Brussels in 2026 using a deliberately less-positive version of the cargo-bike service description. The survey design and MUSA settings were kept unchanged, but the scenario reduced the stated benefits on CO_2, noise, traffic, and safety, while increasing the shipping-cost premium. The changes are summarized in Table 6.

As expected, acceptance declined under this less favorable framing. At the same time, the relative priority structure remained broadly similar. Fig. 5 compares the observed Brussels results with those obtained from the less-positive scenario. The figure shows that acceptance decreases across all criteria, especially for CO_2 emissions and shipping cost, while the ordering of the criterion weights remains largely stable: environmental impact remains the most important criterion, followed by safety, traffic, and noise, whereas shipping cost remains the least important despite its markedly lower acceptance.

This supplementary robustness check should be interpreted cautiously. It is based on a small additional sample from a single city and therefore cannot fully validate the generalizability of the main results. Nevertheless, it provides preliminary support for the view that favorable scenario framing affects the level of acceptance more strongly than the overall ranking of priorities. This reinforces the central interpretation of the paper: the main contribution is not to claim broad support for cargo-bike delivery in general, but to show how a structured multi-criteria analysis can reveal which service dimensions matter most to citizens in different urban contexts, even when the absolute level of acceptance is sensitive to scenario design.

6.2. Reflection on the framework

Compared with a conventional MUSA application, in which all respondents are analyzed together under a single model, the mass-participation multi-city MUSA framework proposed in this study is designed to preserve potentially meaningful heterogeneity across urban contexts. In a standard pooled application, MUSA would be run once on the full sample using the same implementation settings, producing a single aggregate priority structure for all participants. Such an approach is useful when the population can reasonably be treated as homogeneous, but it may conceal important differences when respondents come from cities with distinct urban conditions, mobility systems, and freight-related concerns.

The framework developed here addresses this issue by embedding MUSA within a broader comparative procedure. First, it tests whether responses from different cities can reasonably be treated together. Second, it checks whether within-city subgroup differences are large enough to justify separate modeling. Third, it estimates comparable MUSA models only for the retained city/groups, again using the same implementation settings throughout. In this sense, the framework does not assume a priori that every city must be modeled separately, nor that every city must be split into subgroups. Rather, it provides a systematic procedure (as shown in Fig. 2) for deciding when aggregation is acceptable and when it would mask substantively important differences.

To illustrate the added value of this framework, we also ran a conventional pooled MUSA model on the full sample using the same settings as in the city/group analyses. This pooled model is not used for substantive inference. Instead, it serves as a descriptive benchmark against which the disaggregated framework can be compared. Fig. 6 presents this comparison by displaying the criterion weights for all retained city/groups together with the pooled benchmark. The pooled row shows what the average priority structure would look like if all responses were collapsed into a single MUSA model. In contrast, the disaggregated rows show the criterion weights obtained after applying the framework's grouping logic. The difference between the pooled and disaggregated results illustrates the main advantage of the framework. While the pooled model suggests a broad environmental emphasis, the city/group models reveal the hidden variation that would otherwise remain masked, including the dominant role of noise in Paris, traffic/congestion in Rome, the Urban–Suburban divergence in London, and the especially strong environmental weighting in Warsaw and Brussels. The retained city/group models show that the same hypothetical service is evaluated through different local lenses.

More generally, this suggests that the framework can support future participatory urban-logistics studies in settings where both scale and contextual heterogeneity matter. Its contribution is therefore methodological as much as empirical: it extends a conventional pooled MUSA application into a structured comparative framework that makes aggregation and segmentation decisions explicit rather than arbitrary. The framework also helps clarify how unequal sample sizes should be interpreted in cross-city applications. Because the MUSA models are estimated separately for each retained city/group, larger samples do not mechanically dominate the results of smaller ones. At the same time, sample size still matters for interpretation. Smaller samples may yield less stable or less finely resolved priority structures, whereas very large aggregated samples may smooth over meaningful internal heterogeneity and produce more "average" structures if the underlying population is not first checked for subgroup differences. This is precisely why the framework combines common modeling settings with explicit robustness checks, minimum subgroup-size requirements for segmentation, and a prior validation step on whether groups should be pooled or split. In the present study, the unequal sample sizes partly reflect the practical difficulty of collecting comparable survey data across different cities. This remains a limitation and means that some city-level results should be interpreted with more caution than others.

Table 6
Scenario modifications used in the Brussels sensitivity test. The table compares the original *EcoDeliver* wording with a less-positive version designed to reduce the stated benefits and increase the stated cost premium, while keeping the questionnaire structure and MUSA settings unchanged.

Code	Criterion	Original *EcoDeliver* scenario	Less-positive Brussels sensitivity scenario
g_1	CO_2 emissions	70%–90% lower than diesel vans	20%–40% lower than diesel vans
g_2	Noise	Markedly quieter than vans or mopeds; close to ambient urban noise	Only somewhat quieter; limited overall reduction in perceived street noise
g_3	Traffic	About 75% less road space per delivery than a car	About 25%–40% less road space per delivery than a car
g_4	Safety	Clear reduction in the number and severity of delivery-related accidents	Only limited or moderate safety improvement, depending on street conditions
g_5	Shipping cost	10%–20% higher than standard delivery	20%–35% higher than standard delivery

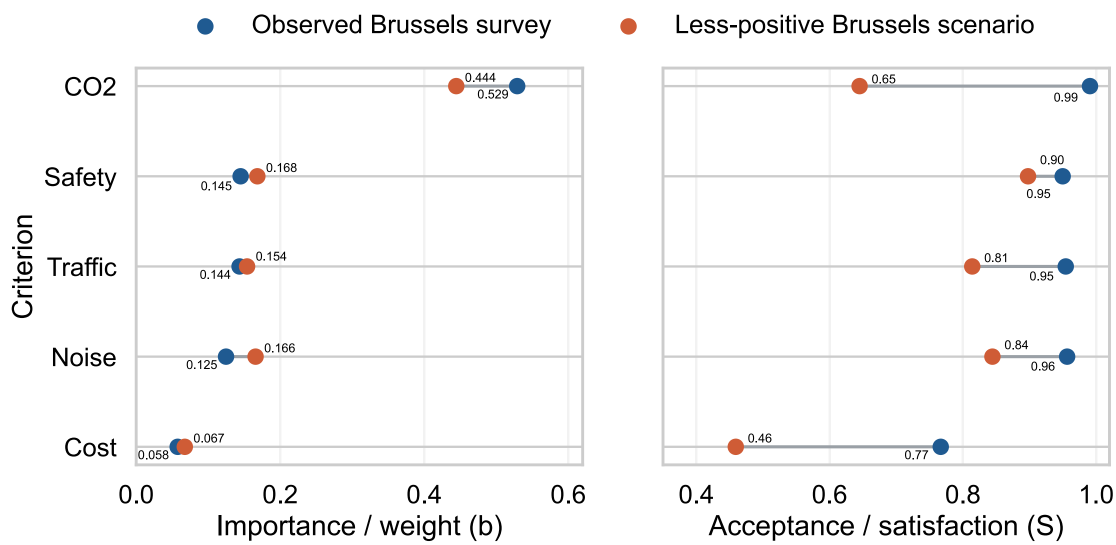


Fig. 5. Brussels sensitivity test comparing the observed survey results with the less-positive scenario. Acceptance levels decline under the less favorable framing, while the relative importance structure remains broadly similar. In both cases, environmental impact remains the most important criterion, whereas shipping cost remains the least important despite its markedly lower acceptance.

At the same time, the framework makes this issue transparent rather than implicit: it can be used not only to compare cities, but also to assess whether differences in sample composition and sample size are large enough to justify aggregation, segmentation, or more cautious interpretation.

7. Conclusion

This paper examined how citizens evaluate a hypothetical cargo-bike last-mile delivery service across six European capital cities using a mass-participation survey and MUSA. The scientific problem addressed in this study is not the technical feasibility of cargo-bike logistics itself, which has already been widely discussed in the urban-freight literature, but how citizens evaluate such a service when it is presented as a multi-attribute urban intervention. In this respect, the study responds to a gap in previous research. Existing work has largely focused either on operational performance, emissions, routing efficiency, and business feasibility, or on citizen perceptions in single-city and mostly descriptive settings. By contrast, the present study combines a harmonized multi-city survey with MUSA to identify not only whether a favorable cargo-bike scenario is positively evaluated, but also which criteria structure that evaluation in different urban contexts.

The results show that the hypothetical *EcoDeliver* scenario receives high acceptance across all retained city/groups. At the same time, the main contribution of the study lies less in the absolute level of acceptance than in the revealed differences in the underlying priority structures. Similar overall evaluations can arise from different configurations of criterion importance. Environmental impact is dominant in several city/groups, Paris is more strongly shaped by noise, Rome by traffic/congestion, and suburban London by safety to a much greater extent than the urban core. In this sense, the novelty of the study lies in moving beyond overall acceptance scores toward a comparative analysis of city-specific priority structures. This is also where the main practical usefulness of the findings lies: they provide guidance on which service dimensions are likely to matter most to citizens when cargo-bike delivery is introduced or expanded in different urban contexts.

From a methodological perspective, the study also demonstrates the usefulness of MUSA in a new application domain. The method makes it possible to work directly with ordinal overall and criterion-level evaluations, recover criterion weights and acceptance value functions, and derive decision-oriented diagnostics such as demanding and improvement indices. Compared with simple descriptive summaries, this yields a more structured account of public acceptance and a more informative basis for comparison across cities. More broadly, the study shows that public acceptance in urban logistics should not be treated as a single undifferentiated construct. Even when overall evaluations are similarly positive, the underlying structure of acceptance may differ substantially across contexts. This is an important theoretical

Fig. 6. Comparative heatmap of criterion importance weights (b_i) across retained city/groups and the pooled benchmark obtained from a single MUSA model applied to the full sample. Darker cells indicate higher relative importance.

implication for comparative urban-freight research, as it suggests that aggregate approval measures alone are insufficient to understand how citizens evaluate logistics innovations.

At the same time, the findings should be interpreted with appropriate caution. First, the study is based on a hypothetical scenario rather than observed experience. Although the scenario was designed to be plausible, respondents evaluated a described service rather than a real one. Moreover, the *EcoDeliver* scenario was framed favorably across several dimensions, which likely contributed to the high acceptance levels. For this reason, the results should not be interpreted as evidence of broad or unconditional public acceptance of cargo-bike delivery in general. Rather, they should be understood as citizens' evaluations of a favorable hypothetical cargo-bike delivery scenario. Second, the samples, while large, are quota-guided online panel samples rather than fully representative samples of each city population. Differences in recruitment difficulty across cities also resulted in unequal sample sizes, which may affect the stability and resolution of some city-level priority structures. Third, the additional Brussels sensitivity check, while useful, is based on a small supplementary sample from a single city and therefore provides only preliminary evidence on the robustness of the identified priority structures under less favorable framing.

The MUSA method itself also entails limitations that should be acknowledged explicitly. Its additive structure implies a compensatory representation of preferences, meaning that weak performance on one criterion may be offset by strengths on others. This may not fully reflect cases in which some respondents hold non-compensatory views, for example if a perceived safety problem would override all other benefits. In addition, the method imposes monotonicity on the learned value functions and derives collective priority structures from stated ordinal evaluations, which necessarily simplify the complexity of real-world attitudes and trade-offs. Thus, while MUSA provides substantial analytical clarity and decision-oriented insight, it remains an approximation of how citizens reason about urban logistics services rather than a complete behavioral model.

Several directions for future research follow from these limitations. One is to compare the present stated-evaluation results with evidence from real-world pilots or implemented cargo-bike delivery schemes. Another is to explore heterogeneity in greater depth using alternative segmentation approaches, such as latent class methods, to identify cross-city respondent profiles that are not captured by geographic grouping alone. Longitudinal studies would also be valuable, as they could examine whether acceptance changes over time as citizens become more familiar with cargo-bike delivery in practice.

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.

Acknowledgments

This work received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101069529, which supported the survey data collection.

Appendix A. Statistical tables

See Tables A.7–A.10.

Table A.7
Median [P25–P75] acceptance by country for each criterion and overall.

Country	CO_2	Noise	Traffic	Safety	Cost	Overall
Brussels	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	4.5 [4.0–5.0]	4.0 [3.0–5.0]	4.0 [4.0–5.0]
Dublin	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	4.0 [4.0–5.0]	4.0 [3.0–5.0]	4.0 [4.0–5.0]
London	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	4.0 [3.0–5.0]	4.0 [4.0–5.0]
Paris	4.0 [4.0–5.0]	5.0 [4.0–5.0]	4.0 [4.0–5.0]	4.0 [4.0–5.0]	4.0 [3.0–5.0]	4.0 [4.0–5.0]
Rome	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	4.0 [4.0–5.0]	4.0 [3.0–4.0]	4.0 [4.0–5.0]
Warsaw	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	5.0 [4.0–5.0]	4.0 [3.0–5.0]	4.0 [4.0–5.0]

Table A.8
Kruskal–Wallis tests across countries for each acceptance variable.

Measure	H	df	p	$\hat{\varepsilon}^2$
Traffic	35.776923	5	1.05e–06	0.012
Safety	24.072388	5	0.00021	0.008
CO_2	13.047740	5	0.0229	0.003
Noise	11.101786	5	0.0494	0.002
Overall	7.940947	5	0.16	0.001
Cost	6.293017	5	0.279	0.001

Table A.9
Significant pairwise country comparisons (Mann–Whitney U, Holm-adjusted p).

Measure	Country A	Country B	p_{adj}	Cliff's δ
CO_2	Paris	Rome	0.017	−0.102
Safety	London	Paris	4.14e–05	0.139
Traffic	London	Paris	2.59e–06	0.153
Traffic	Paris	Rome	7.34e–06	−0.158

Table A.10
Kruskal–Wallis tests of overall acceptance across socio-demographic groups (pooled).

Factor	Levels	H	df	p	$\hat{\varepsilon}^2$
Living	5	115.306006	4	5.37e–24	0.044
Knowledge	4	101.694030	3	6.72e–22	0.039
Shopping	5	35.364290	4	3.91e–07	0.012
Area	2	10.842000	1	0.00102	0.004
Employment	7	18.377595	6	0.00535	0.005
Age	5	7.324180	4	0.12	0.001
Gender	2	1.014079	1	0.314	0.000
Education	4	2.029497	3	0.566	−0.000

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cities.2026.107350.

Data availability

Data will be made available on request.

References

Amaya, J., Delgado-Lindeman, M., Arellana, J., & Allen, J. (2021). Urban freight logistics: What do citizens perceive? *Transportation Research Part E: Logistics and Transportation Review*, *152*, Article 102390.

Ambrosino, G. (2014). *Sustainable urban logistics plan (sulp): The experience of iee enclose project in small and mid-sized European historic towns*. Madrid, Spain.

Bell, M. G. (2021). City logistics and the urban environment. *Urban form and accessibility* (pp. 359–378). Elsevier.

Bjørgen, A., Fossheim, K., & Macharis, C. (2021). How to build stakeholder participation in collaborative urban freight planning. *Cities*, *112*, Article 103149.

Brussels Capital Region (2021). Good move: Regional mobility plan 2020–2030. URL https://data.mobility.brussels/media/publications/plans/prm/goodmove_fr_20210420.pdf. [Accessed 06 April 2026].

Bruyneel, L., Cox, B., Stauffer, A., Vandenthoren, L., Fierens, F., Nawrot, T. S., & Horemans, C. (2025). Positive impact of the introduction of low-emission zones in antwerp and brussels on air quality, socio-economic disparities and health: A quasi-experimental study. *Environment International*, Article 109515.

Caspersen, E., & Navrud, S. (2021). The sharing economy and consumer preferences for environmentally sustainable last mile deliveries. *Transportation Research Part D: Transport and Environment*, *95*, Article 102863.

Chatziioannou, I., Bakogiannis, E., Karolemeas, C., Kourmpa, E., Papadaki, K., & Vlastos, T. (2024). Urban environment's contributory factors for the adoption of cargo bike usage: A systematic literature review. *Future Transportation*, *4*, 92–106.

Chatziioannou, I., Michalakopoulou, K., Tsigdinos, S., Bakogiannis, E., & Nikitas, A. (2026). Evaluating the road environment factors underpinning cargo bike safety and ease of movement: A sustainable urban mobility plan-driven micmac approach. *Transportation Research Part E: Logistics and Transportation Review*, *209*, Article 104756.

Commission, E. E., et al. (2013). A concept for sustainable urban mobility plans. (p. 913). Annex to the Communication from The Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Together towards competitive and resource-efficient urban mobility COM (2013).

Doumpos, M., & Zopounidis, C. (2011). Preference disaggregation and statistical learning for multicriteria decision support: A review. *European Journal of Operational Research*, *209*, 203–214.

Dublin City Council and National Transport Authority (2024). Dublin city centre transport plan 2023: Technical note part 7 – goods movement. URL https://www.dublincity.ie/sites/default/files/2024-04/07-dublin-city-centre-transport-plan-goods-movement-technical-note.pdf. [Accessed 06 April 2026].

European Commission (2021). Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions: The new EU urban mobility framework. Commission Communication COM(2021) 811 final. Directorate-General for Mobility and Transport, Luxembourg. cELEX: 52021DC0811. URL https://op.europa.eu/en/publication-detail/-/publication/ad816b47-8451-11ec-8c40-01aa75ed71a1.

European Commission (2025). Urban transport. URL https://transport.ec.europa.eu/transport-themes/clean-transport-urban-transport/urban-transport_en. [Accessed 03 October 2025].

Foltỳnová, H. B., Vejchodská, E., Rybová, K., & Květoň, V. (2020). Sustainable urban mobility: One definition, different stakeholders' opinions. *Transportation Research Part D: Transport and Environment*, *87*, Article 102465.

França, J. G. d. C. B., Kijewska, K., Iwan, S., Isler, C. A., & de Oliveira, L. K. (2025). Citizens perceptions of urban freight transport: an international comparison. *Transportation Research Interdisciplinary Perspectives*, *29*, Article 101342.

Grigoroudis, E., & Siskos, Y. (2002). Preference disaggregation for measuring and analysing customer satisfaction: The musa method. *European Journal of Operational Research*, *143*, 148–170.

Grigoroudis, E., & Siskos, Y. (2009). Musa: Multicriteria satisfaction analysis. *Customer satisfaction evaluation: methods for measuring and implementing service quality* (pp. 91–121). Springer.

Gunes, S. (2021). *Micro-consolidation practices in urban delivery systems: comparative evaluation of last mile deliveries using e-cargo bikes and microhubs*. University of Washington.

Hagen, T., & Scheel-Kopeinig, S. (2021). Would customers be willing to use an alternative (chargeable) delivery concept for the last mile? *Research in Transportation Business & Management*, *39*, Article 100626.

Hainsch, K., Löffler, K., Burandt, T., Auer, H., Del Granado, P. C., Pisciella, P., & Zwickl-Bernhard, S. (2022). Energy transition scenarios: What policies, societal attitudes, and technology developments will realize the eu green deal? *Energy*, *239*, Article 122067.

Holnicki, P., Kałuszko, A., & Nahorski, Z. (2023). A projection of environmental impact of a low emission zone planned in warsaw, poland. *Sustainability*, *15*, 16260.

Kiba-Janiak, M., Cheba, K., Mucowska, M., de Oliveira, L. K., Piecyk, M., Evangelista, P., Prockl, G., & Rześny-Cieplińska, J. (2024). How to design a sustainable last-mile delivery and returns business model from e-customers' expectations perspective? *Research in Transportation Business & Management*, *56*, Article 101194.

Kiba-Janiak, M., Marcinkowski, J., Jagoda, A., & Skowrońska, A. (2021). Sustainable last mile delivery on e-commerce market in cities from the perspective of various stakeholders. Literature review. *Sustainable Cities and Society*, *71*, Article 102984.

Lenz, B., & Riehle, E. (2013). Bikes for urban freight? Experience in europe. *Transportation Research Record*, *2379*, 39–45.

Llorca, C., & Moeckel, R. (2021). Assesment of the potential of cargo bikes and electrification for last-mile parcel delivery by means of simulation of urban freight flows. *European Transport Research Review*, *13*, 33.

Mayor of London (2018). Mayor's transport strategy. URL https://www.london.gov.uk/sites/default/files/mayors-transport-strategy-2018.pdf. [Accessed 06 April 2026].

Transport for London (2023). Cargo bike action plan. URL https://content.tfl.gov.uk/tfl-cargo-bike-action-plan-2023-acc.pdf. [Accessed 06 April 2026].

Lozzi, R., Marcucci, E., & Gatta, V. (2025). Urban logistics facilities and storytelling. stakeholder engagement, participatory policy-planning and co-creation. *Transportation Planning and Technology*, 1–23.

Maltese, I., Marcucci, E., Gatta, V., Sciullo, A., & Rye, T. (2023). Challenges for public participation in sustainable urban logistics planning: The experience of rome. *Public participation in transport in times of change* (pp. 77–95). Emerald Publishing Limited.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50–60.

Michalakopoulou, K., Vann Yaroson, E., & Chatziioannou, I. (2025). Decoding cargo bikes' potential to be a sustainable last-mile delivery mode: An operations management perspective. *Transportation Planning and Technology*, *48*, 712–734.

Mladenovič, L., Plevnik, A., & Rye, T. (2022). Implementing national support programmes for sustainable urban mobility plans in a multilevel governance context. *Case Studies on Transport Policy*, *10*, 1686–1694.

Montero-Vega, M., Estrada, M., & Thigulla, D. T. (2025). Understanding the impact of socio-demographic differences on acceptance of last-mile delivery technologies: A comparative analysis. *Journal of Transport Geography*, *126*, Article 104244.

Natterer, E. S., Loder, A., & Bogenberger, K. (2026). Effects of paris' cycling policies on vehicular flow: An empirical analysis. *Transportation Research Record*, *2680*, 138–152.

Parajuli, A., & Pojani, D. (2018). Barriers to the pedestrianization of city centres: Perspectives from the global north and the global south. *Journal of Urban Design*, *23*, 142–160.

Philips, I., Cairns, S., De Sejournet, A., Anable, J., Azzouz, L., Behrendt, F., Brand, C., Cass, N., Darking, M., Glachant, C., et al. (2025). E-cargo bikes as a personal transport mode in the uk: Insights from surveys and suburban trials. *Journal of Cycling and Micromobility Research*, Article 100093.

Robichet, A., Nierat, P., & Combes, F. (2022). First and last miles by cargo bikes: Ecological commitment or economically feasible? the case of a parcel service company in paris. *Transportation Research Record*, *2676*, 269–278.

Roma Servizi per la Mobilità (2024). Azioni per lo sviluppo della logistica urbana in attuazione del pums. URL https://romamobilita.it/wp-content/uploads/2024/04/Azioni_sviluppo_logistica_urbana_PUMS_compressed_0.pdf. [Accessed 06 April 2026].

Roman, M., & Fellnhofer, K. (2022). Facilitating the participation of civil society in regional planning: Implementing quadruple helix model in finnish regions. *Land Use Policy*, *112*, Article 105864.

Rudolph, C., & Gruber, J. (2017). Cargo cycles in commercial transport: Potentials, constraints, and recommendations. *Research in Transportation Business & Management*, *24*, 26–36.

Smeeton, N., Spencer, N. H., & Sprent, P. (2025). *Applied nonparametric statistical methods*. CRC Press.

Transport for London (2019). Freight and servicing action plan. URL https://www.london.gov.uk/sites/default/files/tr_18_freight-servicing-action-plan.pdf. [Accessed 06 April 2026].

Ville de Paris (2021). Plan vélo 2021–2026. URL https://cdn.paris.fr/paris/2021/12/08/2fc9cb8ad6db58b6bfde3e6ccfc4c48c.pdf. [Accessed 06 April 2026].

Voll, P., Jennings, M., Hennen, M., Shah, N., & Bardow, A. (2015). The optimum is not enough: A near-optimal solution paradigm for energy systems synthesis. *Energy*, *82*, 446–456. http://dx.doi.org/10.1016/j.energy.2015.01.055.

Warsaw Functional Area (2023). Sustainable mobility plan for the warsaw functional area. URL https://omw.um.warszawa.pl/en/plan-zrownowazonej-mobilnosci-dla-warszawskiego-obszaru-funkcjonalnego/. [Accessed 06 April 2026].

Yang, X., McCoy, E., Hough, K., & de Nazelle, A. (2022). Evaluation of low traffic neighbourhood (ltn) impacts on no2 and traffic. *Transportation Research Part D: Transport and Environment*, *113*, Article 103536.