Transportation Research Part A 210 (2026) 105064

Contents lists available at ScienceDirect

Transportation Research Part A

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

Drivers of trip satisfaction over time: A case study across rapid transit systems in Montreal, Canada

Thiago Carvalho, Ahmed El-Geneidy *

School of Urban Planning, McGill University, Canada

ARTICLE INFO

Keywords:
Trip satisfaction
Metro
Light rail
Bus rapid transit
Longitudinal analysis

ABSTRACT

As public transit agencies struggle to win back riders lost during the COVID-19 pandemic, understanding what drives user satisfaction has become critical to retention and ridership recovery. While most research on transit satisfaction relies on cross-sectional data, this paper adopts a longitudinal approach, integrating the hierarchy of transit needs framework to assess whether satisfaction determinants shift over time and across modes using Montreal, Canada, as a case study. Drawing on panel data from the Montreal Mobility Survey, the analysis focuses on 354 rapid transit commuters surveyed in 2023 and 2024, including users of the metro and the newly launched Bus Rapid Transit (BRT) and Light Rail Transit (LRT) systems. We estimated mode-specific weighted Bayesian models and compared the relative importance of different factors across hierarchical layers. Results indicate that satisfaction with the metro remains stable and increasingly driven by comfort, while the BRT shows improving satisfaction over time with strong reliance on functional factors such as waiting time. For the LRT, both functional and hedonic attributes play central roles in satisfaction, particularly travel times and information availability. This research highlights the importance of sequencing investments to match service maturity, prioritizing foundational needs, such as service reliability, before higher-order features, and investing early in perception-building strategies for new systems. These results provide empirical evidence on the temporal dynamics of satisfaction and offer guidance for transit agencies on strategies to retain current riders and attract new ones in a post-pandemic context.

1. Introduction

Typically, the drivers of trip satisfaction are examined in the context of cross-sectional studies (van Lierop et al., 2018), treating their influence on users as relatively static. Although longitudinal studies are still rare, the literature has shifted toward recognizing these drivers as dynamic, with their relative importance changing as user needs evolve and as transit systems develop and mature (Cao & Cao, 2017; Wan et al., 2016; Zheng et al., 2021). The hierarchy of transit needs (Allen et al., 2019b) provides a useful lens for conceptualizing this evolution, framing satisfaction determinants as a structured set of layers, where higher-level considerations build upon the consistent fulfilment of more basic needs. At the foundational level lie functional attributes such as frequency, reliability, accessibility, and speed; above these are security-oriented factors; and at the top are hedonic attributes such as perceptions of comfort beyond a basic level, payment structures, and user information (Fig. 1). Importantly, this hierarchy is not strictly chronological as older modes can remain at lower-order needs if foundational issues persist, while new systems may start operations meeting higher-order

* Corresponding author.
E-mail addresses: thiago.carvalhodosreissilveira@mail.mcgill.ca (T. Carvalho), ahmed.elgeneidy@mcgill.ca (A. El-Geneidy).

https://doi.org/10.1016/j.tra.2026.105064
Received 24 October 2025; Received in revised form 11 April 2026; Accepted 9 May 2026
Available online 15 May 2026
0965-8564/

expectations.

Understanding the factors that drive user satisfaction is especially critical in the current context, as transit agencies face the dual challenge of retaining existing riders and regaining those lost during the COVID-19 pandemic. The sharp decline in ridership during this period triggered substantial reductions in fare revenue, resulting in significant budget deficits for many agencies (Parker et al., 2021; Qi et al., 2023). Recovery has been slow and uneven, hindered by service cuts and lasting shifts in travel behavior (Haider et al., 2023; Zhao & Gao, 2022). The widespread adoption of telecommuting and other remote activities has disrupted long-standing travel habits (Carvalho & El-Geneidy, 2024; Mohammadi et al., 2023), while many former riders, particularly those with access to a private vehicle or who purchased one during the pandemic, are less likely to have returned to transit (Palm et al., 2022). In this environment, understanding which service attributes influence trip satisfaction can help agencies prioritize improvements that not only meet current riders needs but rebuild trust and attract former users back to the system.

This paper examines temporal shifts in satisfaction with public transit commute trips, with a particular focus on newly established rapid transit services. We investigate changes in the drivers of trip satisfaction between 2023 and 2024 among a panel of 354 rapid public transit commuters, particularly those using Montreal's metro and the newly launched Pie-IX BRT and light rail (LRT) system known as the Réseau express métropolitain (REM). While a one-year period is relatively short for detecting major changes in user perceptions, Montreal offers a unique setting in which multiple rapid transit modes coexist at various stages of operational maturity. The metro is a long-established, high-capacity system, while the Pie-IX BRT opened in November 2022 and the REM began operation on its first branch in July 2023. In this paper, we use maturity in a descriptive sense (established vs early or consolidating operations) and interpret the findings as differences across systems in Montreal, rather than as a definitive test of maturity within a single mode.

We frame our results within the hierarchy of transit needs, using this framework to interpret observed changes, while highlighting dynamics that extend beyond the theory's assumptions. This study addresses the following research questions: (i) how the drivers of trip satisfaction have changed over a one-year period for Montreal's rapid transit services; (ii) do these drivers differ across transit systems with different levels of operational maturity; and (iii) to what extent do the observed patterns align with the hierarchy of transit needs framework.

To analyze these dynamics, we estimate a series of weighted ordered probit models. First, we develop mode-specific models to identify the strongest drivers for each rapid transit service, allowing for comparison across systems that differ in their operational maturity. To assess the empirical alignment of our findings with the hierarchy of transit needs, we investigate the relative importance of attribute-level and hierarchy layer-level factors. This approach contributes to the literature by offering a longitudinal multi-modal analysis of trip satisfaction determinants. The study offers practical insights for transit agencies and planners seeking to prioritize service improvements aiming to retain and attract riders in a post-pandemic context.

2. Literature review

2.1. Understanding the dynamics of trip satisfaction

Research on the drivers of trip satisfaction has been developed based on several conceptual backgrounds. Several studies emphasize service quality frameworks, adapting the SERVQUAL (Parasuraman et al., 1985; Parasuraman et al., 1988) model to assess transit

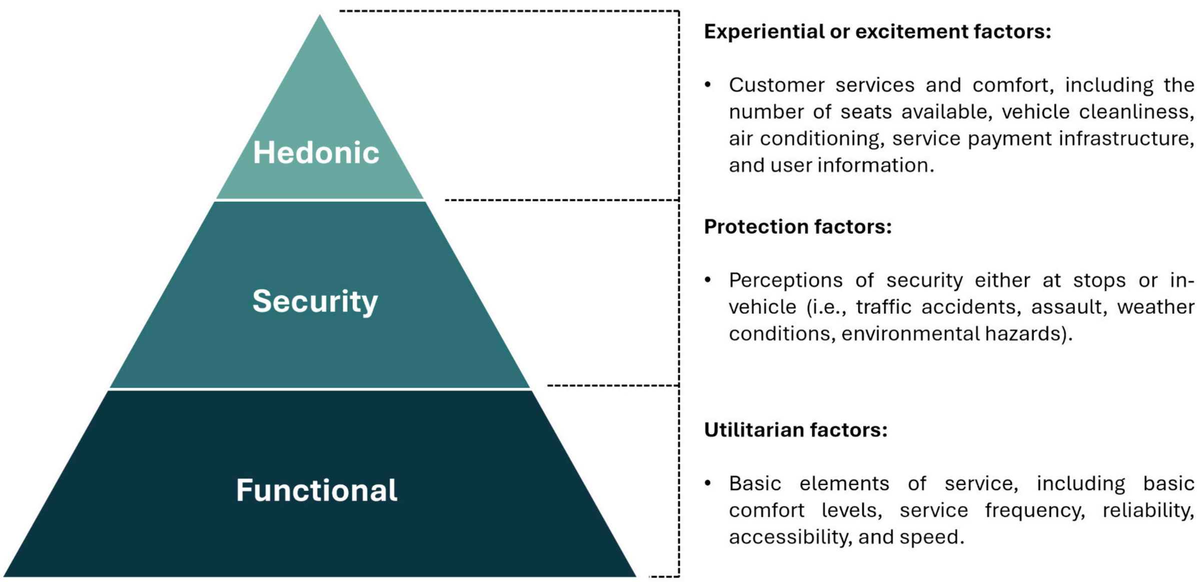

Fig. 1. The hierarchy of transit needs adapted from Allen et al. (2019b).

systems through infrastructure (tangibles and reliability) and experiential factors (responsiveness, assurance, empathy). Others have applied the three-factor theory, which distinguishes between basic factors (or dissatisfiers), performance factors (negative performance has a stronger effect), and excitement factors (positive performance has a stronger effect) (Anderson & Mittal, 2000; Deng et al., 2008; Matzler et al., 2004; Oliver, 1997). Satisfaction has been interpreted as a comparison of expected and perceived performance (i.e., expectation–disconfirmation theory) (Oliver, 1976), indicating that satisfaction would depend not only on service attributes themselves but on how they align with evolving user expectations.

The hierarchy of transit needs (Allen et al., 2019b) builds upon these earlier theoretical approaches. While influential, evidence on its validity remains mixed. The following subsections review findings across BRT, LRT, and metro systems. While these studies were not conducted explicitly through the lens of the hierarchy of transit needs, their results point to patterns that can be interpreted in relation to its layered structure, i.e., support for conditional relevance, evidence of simultaneous needs, or cases where single concerns dominate.

2.1.1. BRT systems

Studies of BRT systems highlight a wide range of factors shaping satisfaction, often reflecting differences in design quality, operational maturity, and local context. Because BRT networks vary from gold-standard segregated corridors to mixed-traffic operations, they provide a useful basis for examining how user evaluations shift as systems evolve. In mature systems, satisfaction drivers often involve factors beyond functional performance. In Bogotá (Colombia), Rodriguez-Valencia et al. (2022) show that mixed-traffic services are judged primarily on speed, reliability, and information, while on the more reliable segregated trunk routes, comfort becomes the dominant concern. Likewise, in Guangzhou's (China) BRT, Cao et al. (2016) report that ease of use, safety, and comfort outrank frequency, reliability, and travel time. A follow-up study (Cao & Cao, 2017) frames comfort as an "excitement" factor, growing in value once basic needs are met.

By contrast, research on newer or still-maturing BRT systems emphasizes foundational attributes. In Amman (Jordan), Al-haj et al. (2024) identify accessibility and operational factors as the strongest drivers, with perceived safety playing a secondary role and hedonic elements such as comfort and information largely absent. Similarly, Chepuri et al. (2024) find that satisfaction in Surat's (India) new BRT system is anchored in reliability and frequency rather than in experiential factors. These findings suggest that when systems are still consolidating, riders focus more heavily on foundational factors.

However, not all results follow this pattern. In Barranquilla's Transmetro, Calvo and Ferrer (2018) report accessibility (functional) and comfort (hedonic) exerting similar influence on trip satisfaction, while Diaz et al. (2023) find that low-income users place greater emphasis on frequency, connectivity, accessibility, and safety, with younger riders in particular evaluating the service more critically. Across diverse contexts, mixed bundles of trip attributes from different hierarchy levels are found to shape satisfaction simultaneously (Saleem et al., 2023; Saxena et al., 2024; Singh & Kathuria, 2023; Zhang et al., 2019).

Overall, BRT studies indicate that satisfaction drivers tend to shift as systems mature, with functional basics dominating in newer systems and experiential concerns gaining weight in established ones. Yet across contexts, multiple attributes often remain important in parallel, suggesting that progression toward higher-order needs does not eliminate the relevance of more basic ones.

2.1.2. LRT systems

Unlike BRT systems, where system maturity appears to structure the relative salience of different attributes, LRT studies frequently highlight parallel influences across multiple dimensions. In Seville (Spain), De Oña et al. (2018) report that service availability (i.e., proximity to stations, waiting time, punctuality) and accessibility (i.e., connection to other modes, universal accessibility, easy to use ticket vendor machines and ticket validators) exert the greatest direct influence on trip satisfaction, reflecting a mix of functional and hedonic concerns. Similar findings are reported in related studies in the same context (De Oña et al., 2016; De Oña et al., 2015; Díez-Mesa et al., 2018), which consistently identify both operational reliability and experiential factors as central to user evaluations. In Eskişehir's EsTram (Turkey), trip satisfaction reflects a bundle of experiential attributes coupled with network operational performance (Yilmaz et al., 2021).

Findings from newer LRT systems underscores the coexistence of needs. In Addis Ababa (Ethiopia), for example, satisfaction is shaped simultaneously by functional drivers (reliable headways, accessibility), security concerns (safe environments), and hedonic attributes, such as clear ticketing systems, information provision, and comfortable spaces. This pattern suggests that even in relatively young systems, users evaluate services based on a combination of operational basics and experiential qualities. Taken together, LRT studies suggest that satisfaction is typically shaped by bundles of functional and hedonic needs, highlighting overlap rather than a clear shift from one level to another.

2.1.3. Metro systems

Metros have been studied extensively across diverse contexts. Evidence on drivers of metro satisfaction is mixed, with some studies supporting the hierarchy of transit needs and others highlighting contradictions. Several studies show that once functional reliability is secured, higher-order factors such as comfort, cleanliness, and information quality become more influential in shaping satisfaction. For instance, research in cities with stable metro networks highlights the growing importance of hedonic factors once on-time performance and reliability are strong (Arancibia et al., 2025; Cao & Cao, 2017; Cao et al., 2016; Nassereddine & Eskandari, 2017). At the same time, functional attributes like reliability and punctuality remain positively associated with satisfaction across different contexts (Allen et al., 2019a; Luo & He, 2021; Osorio-Arjona et al., 2021).

Meanwhile, metro research highlights the persistence of salient unmet needs. Safety concerns and crowding, for example, can dominate user evaluations even in systems with otherwise reliable operations. In Los Angeles, dissatisfaction is closely tied to perceptions of crime and intimidating environments, often linked to homelessness (Shin, 2023). Ouali et al. (2020), in a comparative study of twenty-five systems, found that higher counts of violent incidents significantly reduce satisfaction regardless of service reliability. Similarly, during the COVID-19 pandemic, crowding emerged as a dominant concern (Soltanpour et al., 2020) beyond its usual salience in crowded routes (Machado et al., 2024; Soza-Parra et al., 2019). Many metro studies point to the simultaneous salience of multiple needs. Satisfaction has been found to depend on both operational reliability and experiential qualities such as comfort and cleanliness, as well as security-related concerns (Allen et al., 2020; Efthymiou et al., 2014; Huan et al., 2022; Ibrahim et al., 2025; Li et al., 2022; Saeidi et al., 2020).

Overall, the evidence suggests that metro satisfaction reflects the coexistence of multiple needs. While some studies support the hierarchy by showing a shift toward comfort and information once reliability is secured, others reveal that safety and crowding remain core concerns over time. Taken together, metro research indicates that attributes from different levels of the hierarchy often remain salient simultaneously, with their relative weight shaped by local context and service conditions.

2.1.4. Longitudinal studies

Evidence across BRT, LRT, and metro systems suggests that while satisfaction determinants often shift as systems develop, attributes from multiple layers frequently remain salient in parallel. These mixed findings underline the complexity of applying a hierarchical model to transit satisfaction and point to the need for longitudinal evidence. Such studies can trace how satisfaction drivers evolve over time and clarify whether observed patterns reflect progression through different needs or the continued coexistence of multiple influences.

Although the few longitudinal studies available point to gradual shifts in user priorities, panel evidence remains scarce. Dou et al. (2024), examining Shanghai's (China) metro via a thematic analysis of online reviews from 2010 to 2022, found a shift in service evaluation where user concerns gradually moved away from core operations toward experience-oriented issues like design, staff behavior, and safety in line with the hierarchy's assumptions. Similarly, before-and after studies of new BRT corridors show that at the point of implementation, satisfaction tends to be driven by functional factors, particularly speed and reliability (Conlon et al., 2001; Zheng et al., 2021). Taken together, these studies suggest that user evaluations can evolve from a focus on operational basics toward higher-order needs, but the pace and extent of this shift remain uncertain.

Building upon this research, our study offers two key contributions to the literature on trip satisfaction by LRT, Metro, and BRT. First, it introduces a panel dataset of rapid transit users that allows for the examination of short-term changes in individual-level satisfaction across different rapid transit modes at varying stages of operational maturity. This design provides an opportunity to capture the temporal dynamics impacting how users evaluate their travel experience, shedding light on whether, and how, satisfaction drivers evolve as riders grow accustomed to new services. Second, we build on the hierarchy of transit needs framework by examining changes in predictive salience of functional, security, and hedonic attributes over time, and whether these shifts align with the theory's assumption of conditional relevance. In doing so, this study not only contributes longitudinal evidence to satisfaction research but investigates whether shifts in functional, security, and hedonic needs follow a sequential logic or occur alongside one another within our study context.

3. Case study

Montreal is the second-largest metropolitan region in Canada, home to over four million people. Its rapid transit network is primarily operated by the *Société de transport de Montréal* (STM), which oversees a four-line metro system and one BRT route. The metro began operation in 1966 with the launch of its first two lines, the green and orange lines. A third line, serving the artificial island built to host Expo 67, opened the following year. Since then, the two original lines have undergone several extensions, with the most recent being the orange line's expansion to the city of Laval, completed in 2007. The blue line, the most recent addition to the network, began service in 1986. Currently, it is undergoing an eastward extension that will connect it to the recently launched Pie-IX BRT corridor by the early 2030 s. As of 2024, the metro network carries 1.1 million riders daily making it North America's third busiest metro network (APTA, 2025).

Beyond the metro, recent transit investments have adopted alternative technologies to meet the region's evolving mobility needs. Located on the east side of the island, the Pie-IX BRT is a $523 million CAD project spanning thirteen kilometers, with seventeen stations in operation and three additional stations under construction at the corridor's southern end. The project includes key design elements commonly associated with BRTs worldwide, including center-running dedicated lanes, transit signal priority, all-door and platform-level boarding, and the use of articulated buses. However, the southern segment still operates in mixed traffic until full construction is completed. First announced in 2009, the corridor officially opened in November 2022 after a four-year construction period. By early 2023, it had reached an average ridership of 30,000 passengers per day.

The region is also developing the REM, a fully automated LRT system designed to expand rapid transit coverage across and beyond the island of Montreal. The REM is a 67-kilometre network comprising three branches, with a total projected cost of $9.4 billion CAD. Its implementation is being conducted in multiple phases. The first segment, connecting downtown Montreal to the suburban South Shore region, began operations in summer 2023. One of the remaining branches is scheduled to open in fall 2025, while the third branch is expected in spring 2026 and the final segment to the Montreal-Trudeau International Airport is expected to be operational by 2027. The project is a public private partnership, where it is designed, built, managed, and operated by CDPQ, Quebec's public pension fund, under the jurisdiction of the regional transit authority (ARTM). Initial ridership on the South Shore branch has reached approximately 24,000 passengers per day (CBC News, 2024), with full-network projections estimating 190,000 daily riders once all branches are active (CDPQ Infra, 2017). Since its launch, however, the REM's first segment has experienced frequent service disruptions due to both operational and weather-related issues, which have negatively affected its perceived reliability among users (CBC News, 2025).

The rapid evolution of Montreal's transit network presents a well-suited setting for examining how trip satisfaction evolves over time. The coexistence of a mature metro system alongside two newer rapid transit services offer a unique opportunity to compare satisfaction dynamics across systems at different stages of implementation. This context enables a temporal assessment of how user perceptions evolve as services stabilize and expand. In this way, Montreal provides a case study to test the hierarchy of transit needs, illustrating how the importance of different service features may shift as transit systems evolve. Fig. 2 displays a representation of Montreal's current rapid transit network.

4. Data sources

This study draws on the Montreal Mobility Survey (MMS), a multi-wave, bilingual, online longitudinal survey administered by the Transportation Research at McGill (TRAM) group across the Greater Montreal region. The MMS includes data related to trip satisfaction, particularly for commute purposes, as well as transit perceptions, travel behavior, and socio-demographic characteristics. Across all waves, a combination of recruitment strategies was employed to ensure a large and diverse sample. These included data collection through a specialized marketing firm (Leger), social media campaigns, flyer distribution, and personalized email invitations, following best practices outlined by Dillman et al. (2014).

To ensure data quality and comparability across survey waves, a consistent cleaning strategy was applied. Respondents were excluded based on short completion times, incomplete responses, or multiple submissions from the same Submissions from the same IP address were kept only if the IP address was repeated twice with different Those who placed a pin representing their home, school, and/or work location outside of the Montreal metropolitan area or in water bodies were removed. Panel responses reporting unreasonable change in height or age between waves were excluded. Full details on the survey instrument and cleaning process are available in Carvalho et al. (2025).

To analyze changes in the drivers of satisfaction over time, a panel sample was extracted from the 2023 and 2024 waves of the MMS. To be retained, respondents had to meet the following criteria: (i) they completed the trip satisfaction section in both waves; (ii) they used the same main mode (defined as the mode used to travel the furthest distance during their trip) in both waves, specifically Metro, Pie-IX BRT, or REM; and (iii) both trips were commute-related. After applying these filters, a final panel of 354 respondents was retained, comprising 161 metro users, 78 BRT users, and 115 REM users. Importantly, the 2023 wave reflects user perceptions approximately one year after the launch of the Pie-IX BRT (November 2022) and about three months after the opening of the REM's

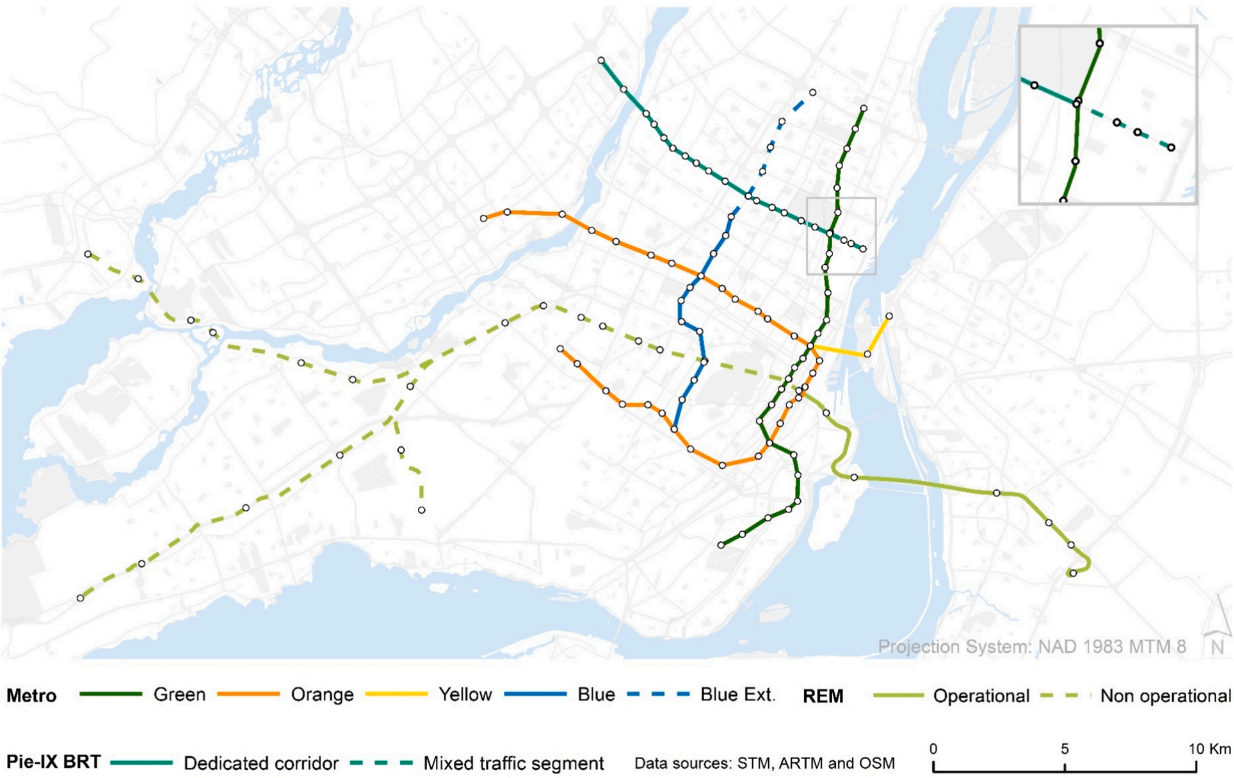

Fig. 2. Montreal's rapid transit network.

first branch (July 2023). The 2024 wave was conducted one year later, capturing more established usage patterns and expectations, particularly for the newer modes. The timeline of data collection is displayed in Fig. 3.

Perceptions of satisfaction can be temporarily influenced by salient events occurring during fieldwork. During the 2023–2024 MMS collection periods, local media reported the following: (i) REM morning-peak service disruptions in early November 2023 (The Canadian Press, 2023); (ii) STM communications in late November 2023 regarding an increased presence of safety personnel in metro stations (City News, 2023); and (iii) an interruption affecting the eastern segment of the Blue line in early October 2024, with replacement bus service and full service restored within approximately one week (Carvalho & El-Geneidy, 2025). Because the MMS does not measure event-specific exposure, we cannot isolate short-term impacts.

Among the mentioned events, the Blue line interruption was the only event with a clearly defined time window, whereas REM disruptions and station safety concerns reflect more persistent conditions; we therefore focus on the Blue line interruption for a sensitivity analysis. In the metro panel (n = 161), only 19 observations fall within the October 4–9, 2024 interruption window, and excluding these observations yields substantively unchanged model estimates, suggesting this event is unlikely to drive the metro results. Localized short-term effects among riders directly exposed to the interruption may nonetheless exist but cannot be identified with our dataset. Moreover, because the interruption was limited to the eastern segment of the Blue line, whose ridership is substantially lower than the Green and Orange lines, any network-wide effect on metro perceptions is also likely to be limited.

4.1. Survey weights strategy

To account for sampling imbalances and reduce sampling biases, survey weights were derived for both waves separately and within each mode-specific analytic sample using the *anesrake* R package (Pasek, 2018), which follows an iterative raking process (DeBell & Krosnick, 2009). The weights were calculated to match the census-tract information of age, income, and gender obtained from Statistics Canada 2021 census (Statistics Canada, 2023), which was retrieved through the *cancensus* R package (von Bergmann et al., 2021). Because publicly available census information does not provide benchmarks for the population of transit riders by mode (Metro/BRT/REM), and because system users may reside beyond a given catchment area, these weights are intended to reduce sociodemographic sampling bias within the study's sample as represented by the census tracts available in the analytic dataset, rather than to exactly reproduce the composition of each system's rider population.

4.2. Weighted descriptive statistics

Table 1 presents weighted descriptive statistics for the 2023 and 2024 panel sample by main mode of transportation as well as the overall sample. Age distribution varies substantially by mode. BRT commuters have a higher proportion of older adults (aged 65 or over), with 18.2% in 2024, compared to 7.9% for Metro and 11.1% for the REM. In contrast, Metro and REM commuters are more likely to fall within the 30 to 44 age group. Gender composition also differs, with a higher proportion of men users among REM commuters (58.9% in 2024) compared to the Metro (45.8%) and the BRT (51.7%). REM commuters are more likely to come from higher-income households: about 30% report annual incomes above $120 k CAD, compared with 24% in the overall 2023–2024 sample. In contrast, over half of BRT users reported annual incomes below $60 k CAD in both years, indicating a higher concentration of low-income riders compared to other modes. Metro users largely mirror the overall sample in terms of household income, with about 40% earning below $60 k CAD, one-third in the middle-income range, and roughly one-quarter above $120 k CAD. This indicates that the Metro serves a broad cross-section of income groups, unlike the BRT or REM which display more skewed distributions.

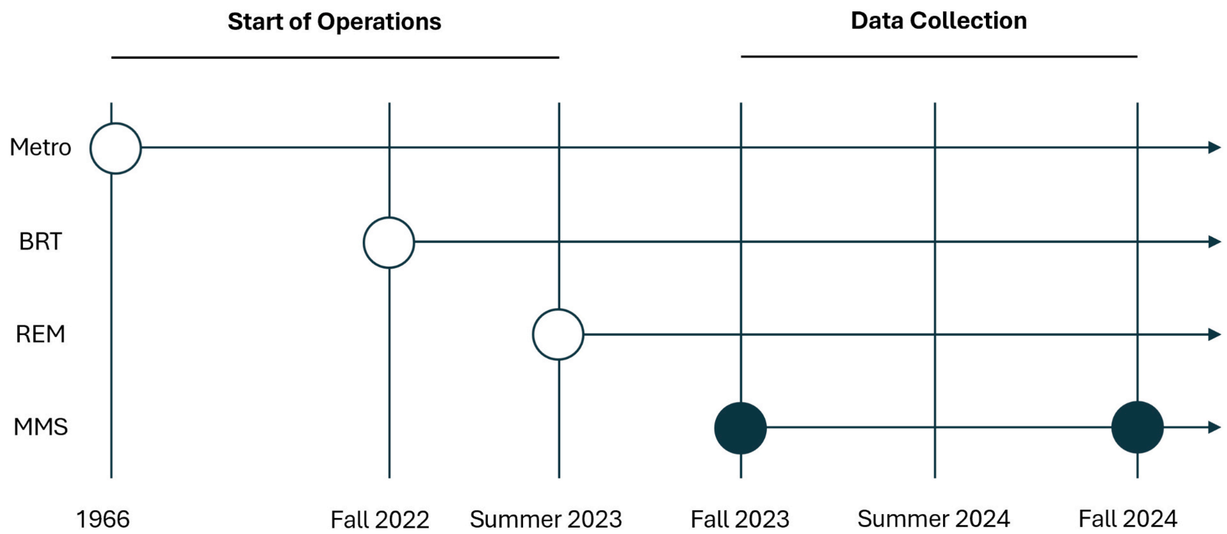

Fig. 3. Timeline of start of operations and data collection.

Table 1
Socio-demographic and transit trip characteristics.

Variable	Metro		Pie-IX BRT		REM		Sample	
	2023	2024	2023	2024	2023	2024	2023	2024
Age								
18–29	35.0%	33.0%	19.9%	16.5%	29.5%	23.0%	29.7%	25.8%
30–44	36.4%	35.8%	44.5%	42.4%	36.5%	33.3%	38.4%	36.8%
45–64	23.6%	23.2%	20.3%	22.9%	25.9%	32.6%	23.5%	25.8%
65+	5.0%	7.9%	15.3%	18.2%	8.1%	11.1%	8.5%	11.5%
Gender								
Men	42.8%	45.8%	48.8%	51.7%	54.4%	58.9%	47.7%	51.1%
Women	57.2%	54.2%	51.2%	48.3%	45.6%	41.1%	52.3%	48.9%
Household income								
< 60 k CAD	39.3%	40.5%	56.6%	55.3%	23.7%	19.5%	38.8%	38.4%
60–120 k CAD	37.0%	34.3%	26.8%	27.2%	45.9%	50.7%	37.2%	37.1%
120 k CAD and over	23.7%	25.2%	16.6%	17.4%	30.4%	29.8%	24.0%	24.5%
Telecommuting								
None	50.6%	49.7%	56.2%	61.5%	43.2%	46.0%	49.8%	51.8%
1–4 days	44.7%	46.6%	32.8%	29.4%	49.5%	50.2%	43.2%	43.1%
5 days	4.8%	3.6%	10.9%	9.1%	7.3%	3.8%	7.0%	5.1%
Car Access	58.8%	51.4%	55.3%	39.6%	80.2%	76.0%	64.3%	55.4%
Travel time (Trip)								
<15 min	3.0%	2.6%	11.3%	9.5%	0.7%	1.1%	4.3%	4.0%
15 to 30 min	14.8%	14.4%	20.7%	15.1%	7.8%	10.1%	14.2%	13.4%
30 to 45 min	36.9%	37.3%	30.1%	33.3%	16.4%	17.6%	29.2%	30.6%
45 to 60 min	22.6%	25.7%	19.8%	22.9%	35.5%	26.6%	25.8%	25.2%
>60 min	22.7%	20.0%	18.0%	19.1%	39.6%	44.5%	26.6%	26.8%
Average distance to nearest station (km)[a]:								
Metro station	2.2 (3.6)	2.3 (3.6)	–	–	–	–	–	–
BRT station	–	–	1.9 (3.5)	1.9 (3.4)	–	–	–	–
REM station	–	–	–	–	5.4 (3.9)	5.2 (3.7)	–	–

[a] Values are formatted as average (standard deviation); Distances are street-network distances from home to the nearest station; they do not imply that respondents walked this distance.

BRT commuters are the least likely to telecommute, with over 60% reporting no telecommuting days in the week before completing the survey. REM and Metro users show higher rates of partial telecommuting (1 to 4 days per week). Car access is much more prevalent among REM users (76% in 2024) than among Metro (51.4%) or BRT commuters (39.6%). Regarding trip duration, REM users tend to have longer commutes. In 2024, over 70% of REM users reported trips lasting more than 45 min, compared to roughly 46% of Metro users and 42% for the BRT. Conversely, Metro and BRT users are more concentrated in the 30 to 45-minute range.

On average, REM users live farther from the nearest station (mean = 5.2 km in 2024; median = 4.2 km) than Metro users (mean = 2.3 km; median = 1.2 km) or BRT users (mean = 1.9 km; median = 0.7 km), highlighting differences in proximity to rapid transit across modes. These are street-network distances from home to the closest station and should not be interpreted as walking access distances, as riders may reach stations via other modes, including feeder bus, cycling, park-and-ride, or drop-off. Moreover, the higher REM distances reflect the South Shore branch's more suburban alignment and wider station spacing. Building on these descriptive insights, we next turn to the modeling strategy used to assess the drivers of trip satisfaction over time.

5. Methods

5.1. Modeling approach

This study focuses on trip satisfaction as the primary outcome of interest. Satisfaction was measured on a four-point Likert scale in response to the following question: *"Overall, I was satisfied with my [mode] experience during this trip,"* where [mode] refers to the respondent's reported main mode of transport, defined as the mode used for the longest distance during the commuting trip. To explore the dynamics of the drivers of trip satisfaction over time, we define a series of Bayesian multilevel ordered probit models at the individual level.

5.1.1. Ordered probit modeling

We select ordered probit modeling because it can reduce Type I and II errors (i.e., detecting non-existing effects or failing to detect existing ones) when data has been collected in a Likert scale compared to converting the scale to a continuous one (Daykin & Moffatt, 2002; Liddell & Kruschke, 2018). Another reason is the increasing use of this technique in the transit satisfaction research (Carvalho & El-Geneidy, 2025; Choi et al., 2021; Fielbaum & Tirachini, 2020; Fu & Juan, 2017; Ouali et al., 2020).

Ordered probit models assume each Likert response comes from an underlying, continuous (unobserved) tendency, such as a person's level of satisfaction (y^*) (Daykin & Moffatt, 2002; Johnson & Albert, 2000). This latent tendency is modeled as the sum of the linear predictor of observed characteristics (η_{it}) and a standard normal error term. For each individual i in year t, latent satisfaction is modeled as follows:

$$\boldsymbol{y}^*_{it} = \boldsymbol{\eta}_{it} + \boldsymbol{\varepsilon}_{it}, \boldsymbol{\varepsilon}_{it} \sim \mathscr{N}(0,1) \tag{1}$$

The observed ordinal (Likert) response (y_{it}) is generated by comparing the latent variable (y^*_{it}) to a set of estimated thresholds (κ). For an outcome with J ordered categories, indexed by $j = 1,\cdots,J$, the model estimates $J-1$ thresholds. For illustration, in the case of three categories, the observed response is defined by two thresholds, κ_1 and κ_2, which define the boundaries between adjacent categories:

$$\boldsymbol{y}_{it} = \begin{cases} 1 if \boldsymbol{y}^*_{it} \le \boldsymbol{\kappa}_1 \\ 2 if \boldsymbol{\kappa}_1 < \boldsymbol{y}^*_{it} \le \boldsymbol{\kappa}_2 \\ 3 if \boldsymbol{y}^*_{it} > \boldsymbol{\kappa}_2 \end{cases} \tag{2}$$

Thus, higher values of the latent variable y^*_{it} make it more likely that the response falls into higher categories. Formally, conditional on the linear predictor η_{it}, the probability that observation y_{it} falls in category j is given by:

$$\boldsymbol{P}(\boldsymbol{y}_{it} = \boldsymbol{j}|\boldsymbol{\eta}_{it}) = \Phi(\boldsymbol{\kappa}_j - \boldsymbol{\eta}_{it}) - \Phi(\boldsymbol{\kappa}_{j-1} - \boldsymbol{\eta}_{it}) \tag{3}$$

In Equation (3), Φ denotes the cumulative distribution function of the standard normal distribution, and κ_{j-1} and κ_j are adjacent thresholds, where $\kappa_0 = -\infty$ and $\kappa_J = +\infty$. The category probabilities represent the probability that the latent satisfaction score falls within the interval defined by adjacent thresholds, given the assumed normal distribution of the error term. In other words, the probability that an observation falls in category j is obtained as the difference between the cumulative probabilities associated with the upper and lower thresholds, evaluated at $\kappa_j - \eta_{it}$ and $\kappa_{j-1} - \eta_{it}$, respectively. Accordingly, when predictors shift y^*_{it} upward (downward), they increase (decrease) the probability of reporting higher values of y_{it}, thereby changing the full set of category probabilities (i.e., the model's implied marginal effects on category probabilities).

The likelihood function represents how plausible the observed responses are under a given set of model parameters. It is constructed by taking the product, across all observations, of the probabilities from Equation (3) corresponding to the categories that were actually observed:

$$\mathscr{L}(\theta) = \prod_{i=1}^{N} \prod_{t=1}^{T_i} P(y_{it}|\eta_{it}, \theta) \tag{4}$$

In Equation (4), $\mathscr{L}(\theta)$ represents the likelihood function, θ the full set of model parameters, i indexes individuals ($i = 1,\cdots,N$), N is the total number of individuals, and t indexes repeated observations for individual i (in this study, two waves: 2023 and 2024). In the Bayesian framework, this likelihood is combined with prior distributions to obtain the posterior distribution of the parameters.

For modeling purposes, the two lowest trip satisfaction categories (i.e., *very dissatisfied* and *dissatisfied*) were combined into a single category labeled *dissatisfied* to comply with the parallel lines assumption. The decision is substantiated by the small share of respondents indicating *very dissatisfied* in both 2023 (5.6%) and 2024 (3.7%), which constrains the possibility of achieving significant statistical differentiation among the dissatisfied categories. A likelihood-ratio test was conducted using a frequentist ordered probit specification to verify the parallel lines assumption. The resulting dependent variable is a three-category ordinal measure of satisfaction, coded as *dissatisfied*, *satisfied*, and *very satisfied*.

5.1.2. Multi-level specification

As observations are repeated for the same individuals over time, responses are not independent. Therefore, we apply a mixed-effects extension of the ordered probit model (Agresti & Natarajan, 2001; Rabe-Hesketh & Skrondal, 2008), in which repeated observations across survey waves are clustered within individuals. This specification allows the latent propensity to vary across individuals via random effects while preserving the ordinal measurement model. We specify a random intercept at the individual level to account for within-person correlation over time and unobserved, time-invariant differences in baseline satisfaction across individuals. The linear predictor η_{it} from Equation (1) can be defined as:

$$\boldsymbol{\eta}_{it} = \boldsymbol{\alpha} + \boldsymbol{u}_i + \boldsymbol{X}_{it}\boldsymbol{\beta}$$

where α is the population-level intercept, X_{it} is a vector of observed covariates for individual i at time t, and β is the corresponding vector of fixed-effect coefficients. The term u_i is a random effect (random intercept) representing individual-specific deviations from the population intercept. In sum, fixed effects describe average relationships shared across individuals, whereas the random intercept allows each individual to have their own baseline level of latent satisfaction. Two model metrics are reported: $\tau 00$ and the intra-class correlation (ICC). $\tau 00$ is the variance of the individual random intercept, reflecting the extent to which individuals differ in their average propensity to report higher satisfaction levels. The ICC indicates the proportion of latent variance attributable to between-individual differences.

5.1.3. Bayesian estimation and interpretation

Bayesian estimation was selected due to its advantages in our modeling context, including greater flexibility to estimate models with complex interactions and multi-level structures, improved handling of small samples, and richer uncertainty quantification through full posterior distributions (Baldwin & Fellingham, 2013). In contrast to frequentist methods, Bayesian models are less prone to convergence issues in small samples and do not rely on asymptotic assumptions for inference, making this approach particularly well-suited for our dataset.

Models were estimated in R using the *brm* function from the *brms* package (Bürkner, 2021), with a cumulative probit link. Survey weights (derived through iterative raking on age, gender, and income) were incorporated to reduce sampling biases. Bayesian estimation relies on Markov chain Monte Carlo (MCMC), which generates a chain of draws that approximate the posterior distribution after warm-up (Johnson & Albert, 2000). In our model estimation, we use four Markov chains with 4,000 iterations each (2,000 warm-up), and a target acceptance probability of 0.95. The target acceptance probability controls how often proposed simulation steps are accepted; higher values generally trade speed for stability. Random intercepts were specified at the individual level to account for within-person correlation over time.

In Bayesian inference, priors are probability distributions assumed before observing the data. The observed data then update these priors through the model likelihood to produce posterior distributions, which summarize what parameter values are most plausible after accounting for both prior information and the evidence in the data. In our modeling, we apply weakly informative priors centered at zero to stabilize estimation and reduce overfitting, particularly in multi-level models, without unduly constraining the results (i.e., posterior distributions). Specifically, we use normal priors for fixed effects $\mathscr{N}(0,2)$, normal priors for thresholds $\mathscr{N}(0,5)$, and an exponential prior for the random-intercept standard deviation $\text{Exponential}(1)$.

Convergence was assessed using the potential scale reduction factor ($\widehat{R}$) and effective sample sizes (ESS). The $\widehat{R}$ statistic compares within-chain and between-chain variance indicating whether independent Markov chains have converged to the same posterior distribution, where values close to 1.00 suggest adequate convergence (Gelman et al., 2013). The effective sample size reflects the amount of independent information in the posterior draws after accounting for autocorrelation in the chains, with larger values indicating more reliable estimation of posterior summaries (Gelman et al., 2013). Convergence diagnostics were obtained from the fitted models and summarized using standard Bayesian workflows in R, including the *bayestestR* package (Makowski et al., 2019).

Model adequacy was evaluated using posterior predictive checks based on the outcome mean: we compared the observed mean to the distribution of means computed from replicated datasets drawn from the posterior predictive distribution, overall and by survey wave. As a complementary summary of model fit, we report marginal and conditional pseudo-R^2 values, which summarize variance explained by fixed effects only and by fixed plus random effects, respectively, on the latent scale.

We summarize posterior distributions using posterior medians and 95% credible intervals (CrI). A 95% CrI can be interpreted as the range within which the model coefficient lies with 95% posterior probability given the data, model and priors. Posterior probability refers to the probability assigned to model parameters (including coefficients) after updating the priors with the observed data. Coefficients are reported on the probit (latent) scale: positive values indicate that higher predictor values shift the latent propensity toward higher satisfaction categories, increasing the probability of higher observed responses while negative values indicate the opposite. Bayesian inference focuses on posterior uncertainty, summarized through credible intervals, rather than p-values and frequentist confidence intervals. We interpret effects as supported when the posterior distribution is clearly concentrated away from zero (e.g., when the 95% CrI excludes zero).

5.1.4. Mode-specific drivers of trip satisfaction

Preliminary stepwise models were estimated to assess the contribution of temporal, socio-demographic/contextual factors, system-

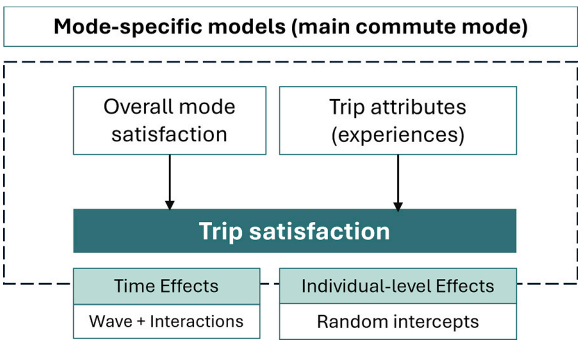

Fig. 4. Structure of the final mode-specific Bayesian multi-level ordered probit models. Trip satisfaction (ordinal) is modeled as a function of overall mode satisfaction and trip-attribute experiences, with wave terms (including interactions) and a respondent-level random intercept.

level perceptions, and trip attributes in explaining trip satisfaction across modes. These exploratory models indicated that trip attributes (e.g., travel time, comfort, information, waiting time) showed the strongest associations with trip satisfaction, while sociodemographic and broader perception variables contributed little once trip attributes were included. Full results from these preliminary models, including predictors that did not show credible directional support, are reported in Appendix A1. Notably, age, gender, income, and most system-level perception variables showed limited directional evidence (i.e., effects were not credibly different from zero), except for overall transit satisfaction, which remained a meaningful predictor. Based on these results, we focus on trip-attribute variables and estimate separate mode-specific models to capture how their effects differ across transit systems and how they evolve over time (Fig. 4). This approach enables the comparison of how drivers of satisfaction vary across systems with distinct levels of operational maturity.

Although travel behavior research suggests that familiarity and habit can reduce deliberative evaluation of service attributes and thus reduce sensitivity to marginal changes in service conditions over time (Gärling & Axhausen, 2003; Gärling & Garvill, 1993), we are not able to directly evaluate whether the importance of different attributes varies with rider experience in this study. Moreover, while overall mode satisfaction was also measured using a Likert scale, it is included in the modeling as an explanatory variable capturing respondents' broader evaluation of the transit system rather than as a second outcome. Our objective is to identify drivers of trip-level satisfaction and their evolution over time, rather than to jointly model multiple satisfaction constructs. Alternative approaches have modeled trip and system satisfaction through a joint ordered mixed logit modeling framework (Henriquez-Jara et al., 2025). In contrast, this study adopts a Bayesian multi-level ordered probit specification that treats overall satisfaction as an explanatory factor, maintaining a parsimonious specification centered on trip-level satisfaction.

5.2. *Examining the hierarchy of transit needs*

Following the identification of predictors with consistent posterior support and the assessment of temporal dynamics, we organize trip attribute variables according to the hierarchy of transit needs (Table 2). Functional attributes include access time, travel time, and waiting time, while hedonic attributes encompass comfort and information. Although security is considered an intermediate level within this framework, safety-related variables were excluded from the analysis because they showed limited posterior support across model iterations (i.e., posterior distributions overlapped zero). Additionally, we introduce a fourth layer, perceived value, reflecting perceptions of cost-benefit, which has been highlighted in the literature as an important factor in determining satisfaction (Fu et al., 2018; Kim & Ulfarsson, 2012; Lai & Chen, 2011; Machado et al., 2018; Sun & Duan, 2019).

To assess how the relative importance of different levels within the hierarchy of transit needs varies across modes and over time, we apply Leave-One-Out Cross-Validation (LOO) to compare model performance when selectively excluding each attribute layer, namely functional (e.g., access time, wait time, and travel time), hedonic (e.g., comfort, information), and value-based (e.g., cost-benefit perceptions) factors. LOO evaluates out-of-sample predictive accuracy by checking how well the model predicts each observation when that observation is left out of the fitting process, summarized using the expected log predictive density (ELPD) (Vehtari et al., 2017). This allows us to evaluate each layer's marginal contribution to predictive accuracy within each mode and survey wave.

LOO was computed using Pareto-smoothed importance sampling (PSIS-LOO) through the *loo* package in R (Vehtari et al., 2017; Vehtari et al., 2025). Rather than refitting the model once for each omitted observation, PSIS-LOO approximates leave-one-out cross-validation from a single fitted model by reweighting posterior draws. Because raw importance weights can become unstable when a small number of draws receive disproportionately large weights, particularly when an omitted observation is highly influential or poorly predicted by the model, Pareto smoothing is applied to the upper tail of the weight distribution. Specifically, the most extreme weights are smoothed using a generalized Pareto distribution, which reduces the influence of extreme values and improves the numerical stability and reliability of the approximation. Therefore, PSIS-LOO approximates leave-one-out cross-validation without requiring repeated model refitting, making it suitable for evaluating out-of-sample predictive performance.

These LOO comparisons are interpreted alongside descriptive statistics of trip-level satisfaction scores. This helps determine the extent to which specific needs are currently being met and whether their fulfillment has changed over time. For instance, if functional needs such as wait time or travel time show marked improvement between waves, it may indicate that basic expectations are increasingly being met. Based on the theory's conditional relevance assumption, we expect that as lower-level functional needs become fulfilled, their predictive importance should diminish. In turn, higher-order needs (i.e., those related to comfort, information, or perceived value) should gain explanatory power in shaping satisfaction. These analyses aim to provide empirical insight into whether transit systems follow the hierarchical progression proposed in theory as they evolve and stabilize over time.

Table 2
Trip-level satisfaction factors by the hierarchy of transit needs.

Hierarchy level	Attribute	Description
Functional	*Access time*	I am satisfied with how long it took me to reach my [mode][a] stop
	Waiting time	The waiting time for [mode] was reasonable
	Travel time	I am satisfied with the length of time I spent on [mode]
Security	*Safety*	I felt safe from crime and unwanted attention when I was on [mode]
Hedonic	*Comfort*	I felt comfortable when I was on [mode]
	Information	Information about [mode] (e.g., schedules, on-board announcements, website, etc.) is easy to understand
Value	*Cost-benefit*	The cost of taking [mode] was reasonable

[a] [mode] reflects the trip main commute mode (i.e., metro, Pie-IX BRT, or REM).

6. Results

6.1. Changes in satisfaction overtime

Fig. 5 reports the temporal dynamics in trip satisfaction between 2023 and 2024 for metro, BRT, and REM commuters. Trip satisfaction among Metro riders remained stable. The share of *satisfied* riders increased marginally by 4.3 percentage points, primarily due to a 4.8 percentage point reduction in *very satisfied* riders. Despite this shift, overall satisfaction remained high. In 2024, 75.1% of riders were *satisfied*, and an additional 17.6% were *very satisfied*. BRT commuters exhibit the most pronounced improvements in trip satisfaction. The share of *satisfied* users increased by 9.4 percentage points, driven mainly by a sharp decline in *dissatisfied* users (-8.5p. p.). These changes increased the overall *satisfied* share to 68.8% in 2024. REM riders' trip satisfaction remained steady between waves, with a 4.9 percentage point increase in *satisfied* users offsetting small reductions in both *dissatisfied* and *very satisfied* shares. These descriptive trends highlight differences in how trip satisfaction has evolved across transit modes. Building on these observations, the next section examines mode-specific drivers of trip satisfaction with multi-level Bayesian models, providing insight into how these factors evolve over time.

6.2. Trip-satisfaction drivers by mode

This section examines how the determinants of trip satisfaction differ across the Metro, BRT, and REM systems using mode-specific models. Each mode-specific model uses the trip attribute variables identified in the preliminary analysis as the strongest predictors of satisfaction. By estimating separate models for each mode, we aim to assess both temporal dynamics and whether different types of infrastructure fulfill user expectations in distinct ways. This approach allows us to evaluate how the salience of trip attributes varies across transit systems at different stages of operational maturity. Table 3 reports posterior medians and 95% credible intervals (CrI). Coefficients are expressed on the probit (latent) scale of an ordered probit model. Positive coefficients indicate that higher predictor values shift the latent propensity toward higher satisfaction, generally increasing the probability of observing higher satisfaction categories (and decreasing the probability of lower categories), while negative coefficients indicate the opposite. For instance, a one-unit increase in a predictor with $\beta = 1.0$ shifts the latent satisfaction propensity upward by one unit, holding other variables constant; the resulting change in category probabilities depends on the thresholds between categories and the individual's baseline position on the latent scale. We interpret an effect as having credible directional support when the 95% CrI does not include zero.

In the estimated ordered probit models, the thresholds define the boundaries between adjacent satisfaction categories on the latent scale (e.g., between "Dissatisfied" and "Satisfied," and between "Satisfied" and "Very satisfied"). They map the latent satisfaction propensity to the observed ordinal responses and are estimated separately for each mode; therefore, they should be interpreted primarily as scale parameters rather than as explanatory effects. In Table 3, the lower threshold (Dissatisfied | Satisfied) is consistently negative, while the upper threshold (Satisfied | Very satisfied) is positive across modes. This reflects the ordering of categories on the latent scale: values below the lower threshold correspond to "Dissatisfied," values between thresholds correspond to "Satisfied," and values above the upper threshold correspond to "Very satisfied.".

Across all models, convergence and sampling efficiency were satisfactory ($\widehat{R}$ values were near 1.00 across parameters; Metro: $\widehat{R} \approx$ 0.9997-1.0004, BRT: $\widehat{R} \approx$ 0.9998-1.0020, REM: $\widehat{R} \approx$ 0.9999-1.0031; ESS values were high, with minimum of approximately 969 and typically in the thousands). Model adequacy was supported by posterior predictive checks based on the outcome mean: the

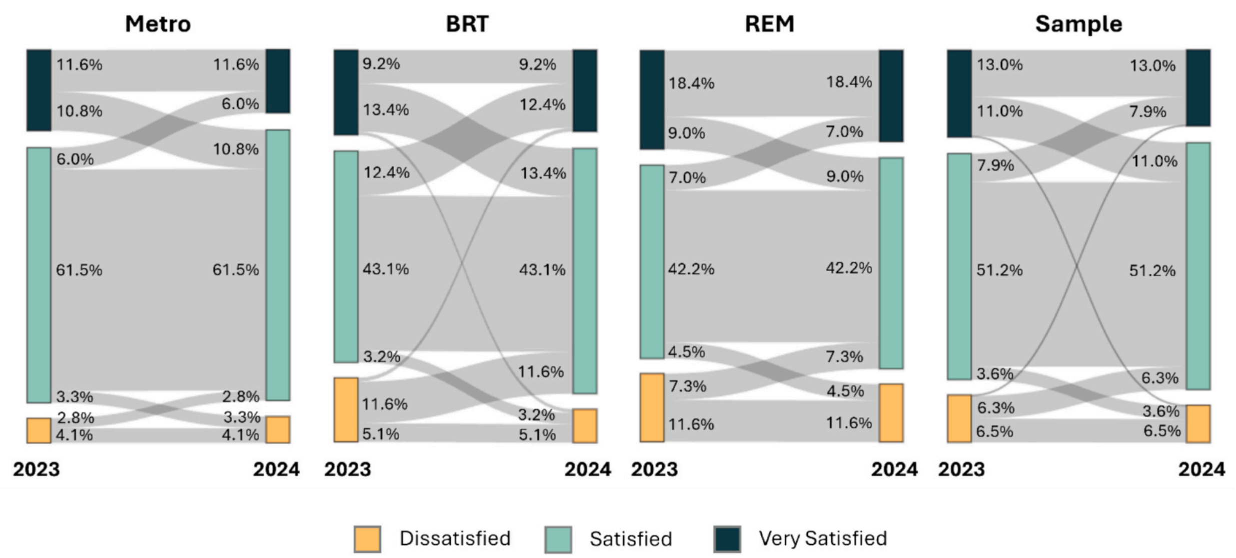

Fig. 5. Trip satisfaction dynamics over time.

Table 3
Mode-specific Bayesian multi-level ordered probit models of trip satisfaction.

	Metro		**BRT**		**REM**	
Predictors	*Median*	*95% CrI*	*Median*	*95% CrI*	*Median*	*95% CrI*
Thresholds						
Dissatisfied | Satisfied	**−6.59**	−10.12 – -4.39	**−3.40**	−5.47 – -1.93	**−4.49**	−7.29 – -2.68
Satisfied | Very Satisfied	**4.46**	2.92 – 6.98	**4.52**	2.71 – 7.40	**3.89**	2.44 – 6.27
Wave *(Baseline: 2023)*						
2024	−0.25	−0.97 – 0.39	**0.88**	0.11 – 1.88	−0.41	−1.20 – 0.25
Overall satisfaction						
Transit satisfaction	**0.83**	**0.32 – 1.53**				
BRT satisfaction			**1.37**	0.57 – 2.63		
REM satisfaction					**1.38**	0.69 – 2.43
Trip attributes (Main mode)						
Access time	0.28	−0.28 – 0.78	0.49	−0.06 – 1.15	**0.55**	0.10 – 1.21
Waiting time	**0.61**	0.12 – 1.22	**1.82**	1.07 – 2.83	0.50	−0.01 – 1.15
2024: Waiting time			**−1.26**	−2.36 – -0.45		
Travel time	**1.35**	0.61 – 2.49	**0.74**	0.09 – 1.61	**1.16**	0.57 – 2.02
Comfort	**1.91**	1.16 – 3.05	0.27	−0.26 – 0.89	**0.67**	0.13 – 1.48
Information	**0.83**	0.28 – 1.62	**0.71**	0.17 – 1.46	**1.03**	0.52 – 1.80
Cost-benefit	**1.15**	0.57 – 2.04	**0.78**	0.11 – 1.63	0.29	−0.19 – 0.77
Random Effects						
τ_{00}	0.22		0.30		0.42	
ICC	0.15		0.09		0.02	
N	161 $_{ID}$		78 $_{ID}$		115 $_{ID}$	
Observations	322		156		230	
Marginal R^2/Conditional R^2	0.645/0.806		0.653/0.766		0.737/0.824	

Note: Median reflects posterior medians (probit-scale coefficients, β); 95% CrI denotes the 95% credible interval. Table 3 reports the final parsimonious mode-specific models; full exploratory stepwise results, including non-credible estimates, are reported in Appendix A1.

observed mean fell within the posterior predictive distribution overall and by wave. As a complementary summary of fit on the latent scale, marginal/conditional pseudo-R^2 values indicate substantial explanatory power from fixed effects alone and additional improvement when including individual random intercepts (Table 3).

To interpret the contribution of the random intercepts, we report random-intercept variance (τ00) and the intra-class correlation (ICC). τ00 indicates how much respondents differ from one another in their baseline tendency to report higher satisfaction (larger values mean greater between-person differences). The ICC (0–1) indicates how similar repeated responses from the same respondent are (higher values mean stronger within-person consistency over time). In our results, Metro shows the highest within-person consistency (ICC = 0.15) and the lowest between-person dispersion (τ00 = 0.22); BRT is intermediate (ICC = 0.09; τ00 = 0.30), whereas REM shows little within-person clustering (ICC = 0.02) and higher between-person dispersion (τ00 = 0.42). These findings suggest that Metro respondents show greater within-person stability across repeated observations, whereas REM shows minimal within-person clustering, implying that variation is primarily within individuals.

The Metro model provides insight into how satisfaction determinants manifest in the most mature of the three systems. Results indicate that satisfaction with the metro is shaped by a combination of hedonic and functional trip attributes, with little evidence of temporal effects observed between 2023 and 2024. This temporal stability may reflect the metro's established infrastructure and service delivery, where changes in satisfaction are likely less driven by evolving operational characteristics. On the probit (latent) scale, comfort shows the strongest positive association with higher satisfaction (β = 1.91, 95% CrI: 1.16 – 3.05), followed by travel time (β = 1.35, 95% CrI: 0.61 – 2.49) and cost-benefit (β = 1.15, 95% CrI: 0.57 – 2.04), ceteris paribus. Information (β = 0.83, 95% CrI: 0.28 – 1.62) and waiting time (β = 0.61, 95% CrI: 0.12 – 1.22) also exhibit positive associations with satisfaction, although to a lesser extent, all else equal.

The BRT model reveals distinct dynamics reflective of a newer and evolving infrastructure. Unlike the Metro, the BRT model shows credible evidence of a temporal shift between 2023 and 2024 (β = 0.88, 95% CrI: 0.11 – 1.88), suggesting that perceptions of the BRT experience are still evolving, keeping all other variables constant. Notably, overall satisfaction with the BRT system itself is strongly associated with trip satisfaction on the probit (latent) scale (β = 1.37, 95% CrI: 0.57 – 2.63), with a larger magnitude than most attribute-specific predictors all else equal. This finding highlights the importance of overall impressions of the new infrastructure during its early operational phase. Among attribute-specific predictors, waiting time was the most influential: in 2023 it is strongly associated with higher trip satisfaction (β = 1.82, 95% CrI: 1.07 – 2.83), but this association is substantially weaker in 2024, as indicated by a negative interaction (2024 × waiting time: β = − 1.26, 95% CrI: −2.36 – -0.45). This sharp reduction may not reflect an improvement in service, as a separate difference-in-means test indicates that satisfaction with waiting time remains low and does not differ between waves (mean = 0.5; p = 0.641). Instead, it may suggest that users have adapted their expectations or grown accustomed to the service frequency, diminishing its impact on satisfaction. Other relevant predictors include travel time (β = 0.74, 95% CrI: 0.09 – 1.61), cost-benefit (β = 0.78, 95% CrI: 0.11 – 1.63), and information (β = 0.71, 95% CrI: 0.17 – 1.46), pointing to a mix of functional and hedonic considerations. In contrast, comfort does not show directional support in the BRT model, possibly reflecting lower expectations or inconsistency in the quality of the BRT experience.

The REM model provides insight into satisfaction determinants for the newest transit infrastructure in the region. Unlike the BRT, the REM model shows limited evidence of a temporal shift between 2023 and 2024, suggesting either that user perceptions have stabilized early or that service conditions have remained unchanged. However, as with the BRT, overall satisfaction with the REM system is strongly associated with trip satisfaction on the probit (latent) scale (β = 1.38, 95% CrI: 0.69 – 2.43), all else equal, underscoring the weight of general perceptions during the early operational phases of a new system. Among trip-level attributes, travel time (β = 1.16, 95% CrI: 0.57 – 2.02) and information quality (β = 1.03, 95% CrI: 0.52 – 1.80) show the strongest positive associations with higher satisfaction, reflecting both functional and hedonic components. Comfort also shows credible directional support (β = 0.67, 95% CrI: 0.13 – 1.48), as does access time (β = 0.55, 95% CrI: 0.10 – 1.21), though with smaller magnitudes. This contrasts with Metro and BRT, where access-time satisfaction is tightly clustered around ‘satisfied,’ limiting its ability to explain variation in overall trip satisfaction once other attributes are included; REM exhibits greater dispersion, allowing access conditions to emerge as a more discriminating predictor. The relevance of access time may stem from the REM’s suburban-oriented development in the South Shore branch, which often requires longer first- and last-mile travel compared to more centrally located modes like the metro. In contrast, waiting time and cost–benefit do not show credible directional support in the REM model.

Overall, travel time emerges as a consistently important determinant of trip satisfaction across all three transit systems, regardless of their operational maturity. This suggests that riders consistently view reasonable travel times as a baseline indicator shaping their overall experiences. Beyond this commonality, the relative salience of other attributes varies across modes. Next, we assess whether these drivers hold at a more aggregate level by analyzing how each hierarchical layer contributes to predictive accuracy.

6.3. Examining the hierarchy of transit needs

This section examines whether the hierarchy of transit needs provides a useful lens for understanding satisfaction dynamics, beginning with descriptive statistics and moving to model-based comparisons. We begin by examining satisfaction levels for individual trip attributes to evaluate whether basic service aspects, such as access, waiting, and travel time, are consistently met across modes and over time. We then examine whether changes in these basic service attributes translate into shifts in the importance of higher-order factors, using model comparisons across the different layers of the hierarchy.

Table 4 presents descriptive statistics of satisfaction with individual trip attributes, focusing on whether lower-level needs, such as access, waiting, and travel time, are fulfilled across modes and years. Metrics reflect the average of responses on a four-point Likert scale (from very dissatisfied to very satisfied, without a neutral option), which was converted to a −2 to 2 range; higher values therefore indicate greater satisfaction. By assessing improvements or stagnation over time, we evaluate the extent to which each system meets riders’ basic service expectations. This evaluation is important because, according to the hierarchy of transit needs, once these basic expectations are met, higher-order attributes should become more salient in shaping satisfaction.

6.3.1. Lower-order needs: Functional factors

At the functional level, satisfaction with access, waiting, and travel time is generally high, though patterns vary by mode. Metro users reported consistently positive ratings across all three attributes, with a statistically significant improvement in satisfaction with access time between 2023 and 2024 (p = 0.007). BRT users also expressed satisfaction with access and travel time, yet waiting time remained a weak point (mean = 0.5), despite the corridor’s high-frequency service. This likely reflects continued exposure to mixed-traffic operations and incomplete infrastructure, particularly along the southern and northern sections of the BRT corridor. While waiting time lost predictive power in 2024, this shift likely reflects user adaptation rather than actual service improvements. For REM users, satisfaction with waiting and travel time was high, consistent with the system’s segregated right-of-way and frequent service. However, access time lagged behind, reflecting longer average distances to stations and limited feeder connectivity.

Overall, the Metro reliably fulfills basic service expectations, while both the BRT and REM continue to face functional gaps, i.e., waiting time for the BRT and access time for the REM. The Metro and REM both benefit from full segregation, which supports

Table 4
Satisfaction with trip-level attributes by mode over time.

Hierarchy level	Variable	Metro		BRT		REM		Sample	
		2023	2024	2023	2024	2023	2024	2023	2024
Functional	*Access time*	0.8 (1.1)	1.1 (0.9)	0.9 (1.0)	0.8 (1.1)	0.3 (1.3)	0.3 (1.4)	0.7 (1.2)	0.8 (1.1)
	Waiting time	1.0 (0.9)	1.1 (0.7)	0.5 (1.3)	0.5 (1.2)	1.1 (1.0)	1.2 (0.8)	0.9 (1.0)	0.9 (0.9)
	Travel time	0.9 (1.0)	0.9 (0.9)	1.0 (1.0)	0.8 (1.1)	0.8 (1.3)	0.9 (1.2)	0.9 (1.1)	0.9 (1.0)
Security	*Safety*	0.7 (1.1)	0.8 (0.9)	0.8 (1.0)	0.8 (1.0)	1.2 (0.8)	1.3 (0.6)	0.9 (1.0)	1.0 (0.9)
Hedonic	*Comfort*	0.9 (1.1)	0.9 (0.9)	0.8 (1.2)	0.4 (1.2)	1.0 (1.1)	1.2 (0.9)	0.9 (1.1)	0.8 (1.0)
	Information	0.9 (1.0)	1.1 (0.8)	0.8 (1.2)	0.9 (1.0)	0.7 (1.2)	1.1 (0.9)	0.8 (1.1)	1.0 (0.9)
Value	*Cost-benefit*	0.5 (1.2)	0.5 (1.1)	0.9 (1.0)	0.7 (1.1)	0.3 (1.3)	0.4 (1.2)	0.6 (1.2)	0.5 (1.1)

reliability and fast service, whereas the BRT's exposure to mixed traffic and incomplete infrastructure contributes to uneven fulfillment of basic needs. Importantly, the REM's access time issues are largely external to its operational design, reflecting longer distances to stations and limited feeder transit, while the BRT's waiting time issues appear more intrinsic to its corridor operations. According to the hierarchy of transit needs, these unresolved challenges suggest that functional attributes should remain the most salient drivers of satisfaction for the BRT and REM, while higher-order factors are more likely to become prominent in the Metro, where foundational needs are already met.

6.3.2. Higher-order needs: Safety and hedonic factors

At higher levels of the hierarchy, attributes related to safety, comfort, information, and cost-benefit show distinct patterns across modes. Safety consistently received high ratings (means $\geq$ 0.8) across all systems and showed no meaningful change over time. Because it did not show a statistically significant influence on trip-level satisfaction in the models, it was excluded from further analysis. Other higher-order attributes, however, varied more noticeably across modes. Hedonic attributes, including comfort and information, were strongly fulfilled in both the Metro and REM. Metro users reported stable, high satisfaction with both attributes, while REM users noted statistically significant gains in comfort ($p = 0.05$) and information ($p = 0.01$), reflecting improving user experiences during the service's first operational year. In contrast, BRT users experienced a notable decline in comfort between waves ($p = 0.03$), though information remained moderately positive. Finally, cost-benefit perceptions were lowest for the REM (mean = 0.4 in 2024), moderate for the Metro (0.5), and highest for the BRT (0.7).

6.3.3. Model-based comparisons

To formally evaluate these patterns, we compare predictive performance across hierarchical layers using Leave-One-Out (LOO) cross-validation (Fig. 6). For the BRT, we report results separately for 2023 and 2024 to capture the temporal variation observed in the models. In contrast, Metro and REM results are presented as single estimates, as the models showed no significant changes between survey waves.

The LOO results largely support theoretical expectations. For the Metro, the largest decline in predictive accuracy occurred when hedonic needs were excluded ($\Delta = -31.9$), followed by value ($\Delta = -18.3$) and functional attributes ($\Delta = -9.8$), indicating that higher-order needs are key drivers of satisfaction in this context. In the REM, the strongest impact came from functional needs ($\Delta = -21.6$), though hedonic needs played a notable role ($\Delta = -11.9$). This could suggest a transitional stage: intrinsic operational factors such as waiting and travel times are well fulfilled, opening space for hedonic attributes to matter, but access remains a persistent gap that keeps functional needs central. The coexistence of strong functional and emerging hedonic influences highlights a partial alignment with the hierarchy of transit needs, though not a complete one, since not all lower-order needs are met.

The BRT case is more nuanced. In 2023, functional needs were the dominant predictors ($\Delta = -21.6$), but this influence diminished in 2024 ($\Delta = -7.0$), even though satisfaction with waiting times remained low. This shift does not appear to reflect improved perceptions of service, but rather a possible de-prioritization of functional concerns, perhaps due to adaptation or shifting expectations. Concurrently, the predictive role of hedonic attributes shows only modest change over time (from $\Delta = 0.25$ to -2.96). While this suggests some increase in the relevance of higher-order needs, the evidence remains limited. Importantly, the attenuation of functional needs is still observed when comfort is excluded from the specification, indicating that this pattern is not driven by the inclusion of comfort but is more closely linked to the reduced salience of waiting time over time. Importantly, such reordering of needs in the absence of improved satisfaction at the foundational level is not anticipated by the theory and may point to user adaptation that warrants further investigation by future research.

The three cases illustrate varying degrees of alignment with the hierarchy of transit needs: the Metro reflects its expectations most closely, the REM appears transitional as both functional and hedonic needs matter, and the BRT diverges as functional concerns remain

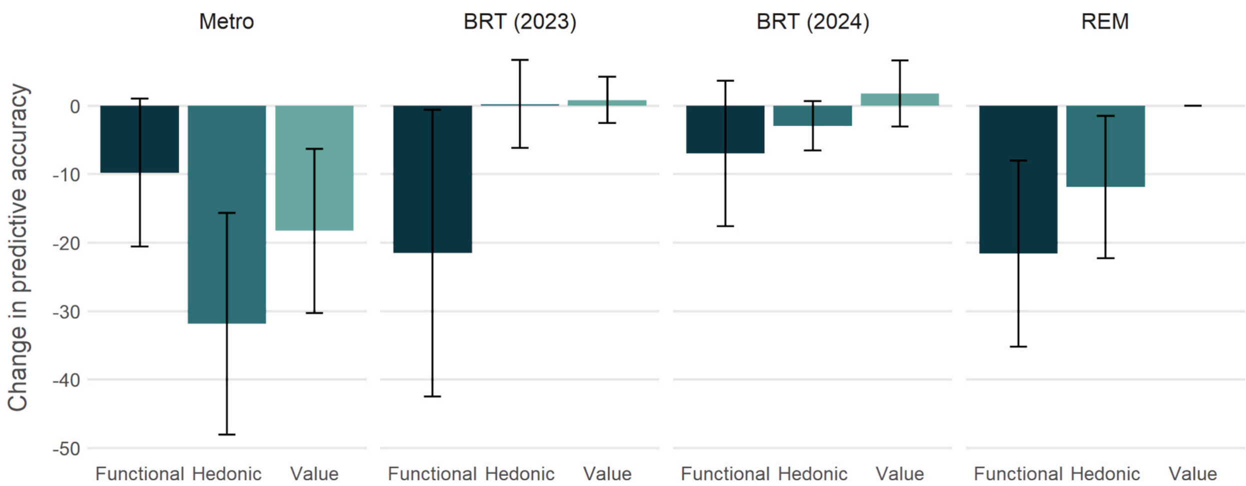

Fig. 6. Changes in predictive accuracy when functional, hedonic, and value needs are excluded from the base models by mode and wave.

unresolved yet appear to lose salience over time.

7. Discussion

7.1. The temporal dynamics of trip satisfaction

Most prior research on public transit satisfaction has relied on cross-sectional data. While informative, such studies cannot assess temporal dynamics, including how satisfaction determinants change (or remain stable) over time, particularly as systems move from early operation toward more mature service conditions. Our study contributes panel evidence from rapid transit commuters in Montreal, Canada, surveyed in 2023 and 2024 across two recently launched BRT and LRT systems and the more established Metro network. Because the panel spans one year, results should be interpreted as short-term dynamics rather than long-term effects. Still, Montreal provides a valuable setting in which multiple modes coexist at distinct stages of maturity, allowing a comparative view of satisfaction drivers across systems.

For the Metro, satisfaction is shaped by a mix of hedonic and functional attributes, with no clear evidence of temporal effects between 2023 and 2024. Our findings suggest that metro users expect a balance of service performance and user experience, with hedonic factors becoming particularly salient, likely because operational reliability is already perceived as stable. This pattern is consistent with cross-sectional studies of established metro networks where comfort, cleanliness, and information gain importance once basic performance needs are met (Arancibia et al., 2025; Cao et al., 2016; Nassereddine & Eskandari, 2017), though functional attributes such as reliability remain important (Allen et al., 2019a; Luo & He, 2021). Longitudinal evidence from Shanghai points to a gradual shift toward hedonic concerns over a 12-year period (Dou et al., 2024). In Montreal, our findings capture a shorter window, showing that hedonic factors are already central alongside functional ones.

For the BRT, the patterns suggest distinct temporal dynamics, consistent with a system still in the process of consolidation. Unlike the Metro, an improvement in overall satisfaction is observed between 2023 and 2024, indicating that user evaluations of the Pie-IX corridor continue to evolve. Satisfaction with the BRT system as a whole exerts a stronger influence on trip satisfaction than most specific attributes, underscoring the weight of general impressions during the early operational phase. Waiting time appears especially salient in 2023 but declined sharply in importance by 2024, despite no measurable improvement in satisfaction with this attribute. This points to riders adjusting their expectations or becoming accustomed to service frequency, reducing its role in shaping evaluations. This interpretation is consistent with expectation-disconfirmation theory (Oliver, 1976), which suggests that satisfaction depends not only on service conditions themselves but also on how they compare with users' evolving expectations. Moreover, this finding is in line with previous research on passengers' travel time perceptions following express route service improvements leading to an initial overestimation of expected travel time savings (Diab & El-Geneidy, 2014). Other predictors such as travel time, cost-benefit, and information contributed meaningfully, while comfort played a more limited role, suggesting lower expectations around higher-order features at this stage.

These findings echo cross-sectional research on new BRT systems in Amman and Surat, where foundational attributes such as reliability and accessibility dominated early evaluations (Al-haj et al., 2024; Chepuri et al., 2024). Our longitudinal findings extend this evidence by showing that the salience of functional attributes, such as waiting times, can diminish over time even in the absence of service improvements. This highlights the importance of expectation adjustment in shaping satisfaction, a process that cross-sectional studies cannot fully capture and that aligns with the expectation-disconfirmation framework.

For the REM, findings indicate stability in satisfaction determinants between 2023 and 2024, with no substantial temporal effects detected. As with the BRT, satisfaction with the system as a whole play a prominent role, underscoring the influence of broad perceptions in the early stages of operation. At the attribute level, trip satisfaction appears to reflect a combination of functional and hedonic dimensions, with travel time and information being particularly salient. The salience of information may relate to the REM's recurrent service disruptions, which are likely to heighten users' attention to real-time updates and operational alerts. Comfort and access time also contribute, with the latter possibly linked to the suburban orientation of the South Shore branch, where longer first- and last-mile trips are common. In contrast, waiting time and cost-benefit appear less central to user evaluations of the service.

These patterns align with findings from other LRT studies, which frequently highlight bundles of functional and hedonic attributes shaping satisfaction in parallel (De Oña et al., 2016; De Oña et al., 2018; De Oña et al., 2015; Díez-Mesa et al., 2018; Yilmaz et al., 2021). Evidence from newer systems, such as Addis Ababa, similarly points to the coexistence of needs, with users evaluating services based on both operational performance and experiential qualities from early operational stages (Efthymiou et al., 2014; Ibrahim et al., 2025). Our findings reinforce this pattern, rather than showing a sequential shift in priorities, where satisfaction appears to reflect a mix of functional, hedonic, and contextual concerns even during the system's first operational year.

In sum, the three mode-specific models point to different trajectories of satisfaction across systems operating at various stages of maturity. In the Metro, determinants remained stable over time, with hedonic concerns salient alongside core functional attributes, consistent with patterns expected in a long-established system. The BRT, by contrast, showed signs of shifting user priorities, with the influence of waiting time declining despite no improvement in performance, which may indicate expectation adjustments as the service becomes familiar. The REM displayed no temporal change but a reliance on both functional and hedonic factors, consistent with evidence from other light-rail systems and possibly indicating a transitional state or concurrent prioritization of reliability and rider experience. Together, these patterns illustrate how user satisfaction with rapid transit can remain stable, adapt in the absence of measurable service improvements, or reflect bundled influences during early consolidation (i.e., functional and hedonic factors). More broadly, the results show how longitudinal analysis adds nuance to cross-sectional findings by capturing different ways satisfaction evolves across modes within the same metropolitan context.

7.2. The hierarchy of transit needs

We explored the extent to which our empirical findings align with (or diverge from) the hierarchy of transit needs framework (Allen et al., 2019b). Rather than testing the theory in a definitive sense, we assess whether the observed satisfaction dynamics across systems at different stages of operational maturity are consistent with its assumptions. This matters because, if empirically supported, the theory's conditional relevance assumption can inform service planning by offering a roadmap for how priorities might change as systems become more established.

Although previous studies provide mixed evidence, two patterns recur in the literature. First, hedonic factors (e.g., comfort, information) tend to become more influential once functional performance is stable. Second, poor performance in any attribute, regardless of its place in the hierarchy, can increase its role in shaping overall satisfaction. Our findings add to this debate by examining both individual satisfaction drivers and the collective contribution of each hierarchical layer. To do this, we compared models that progressively excluded functional, hedonic, or value-related attributes and assessed how much predictive accuracy declined in each case. This approach allows us to evaluate not just whether individual factors matter, but how much weight each broader layer of the hierarchy carries in shaping overall satisfaction.

At the attribute level, findings point to a partial alignment with the hierarchy of transit needs. Metro riders' emphasis on comfort reflects a shift toward higher-order considerations once foundational needs are met. BRT users continued focus on waiting time illustrates how functional gaps dominate evaluations in maturing systems. The REM's reliance on both travel time and information suggests that functional and hedonic concerns can operate simultaneously. Across modes, travel time remains a consistently important determinant. These results indicate that while the hierarchy offers a useful organizing principle, its conditional relevance is not absolute, with functional and hedonic attributes often operating in tandem.

Stronger support emerges when comparing the aggregate contribution of each layer (Fig. 7). In the Metro, removing hedonic attributes from the model led to the largest decrease in predictive accuracy, followed by value and functional attributes supporting the theory's core assumption. For the REM, results indicate a layered but potentially transitional structure during its early operational stage. Functional needs exerted the strongest influence, yet hedonic factors such as information and comfort also played a notable role. This combination suggests that while core operational expectations like waiting and travel times are being met, unresolved challenges with station accessibility keep functional needs central. At the same time, the concurrent influence of hedonic attributes indicates a partial rather than complete alignment with the hierarchy of transit needs.

The BRT results were more complex. In 2023, functional attributes dominated the satisfaction model, reflecting low satisfaction levels with waiting time and aligning with the hierarchy of transit needs. However, by 2024 the explanatory power of functional attributes declined substantially. As previously highlighted, this reduction does not appear to reflect actual improvements in perceived waiting time, but rather a potential de-prioritization of these concerns. While this shift could suggest that higher-order needs are becoming more relevant, the corresponding change in hedonic contribution remains modest and does not clearly indicate what, if anything, is replacing functional concerns in users' evaluative framework. Importantly, the attenuation of functional needs is also observed when comfort is excluded from the specification, suggesting that this pattern is not driven by the inclusion of comfort but is more closely linked to the reduced salience of waiting time over time.

Overall, the findings suggest that the hierarchy of transit needs serves best as a broad organizing principle rather than a strict rule. At the attribute level, functional and hedonic drivers often operate in tandem, particularly in systems like the REM where some foundational gaps persist alongside strong operational performance. At the aggregate layer level, clearer patterns emerge: the metro aligns most closely with the theory's expectations, the REM illustrates a potential transitional stage, and the BRT diverges as functional concerns appear to lose salience despite being unresolved. This layered perspective highlights both the value and the limits of the

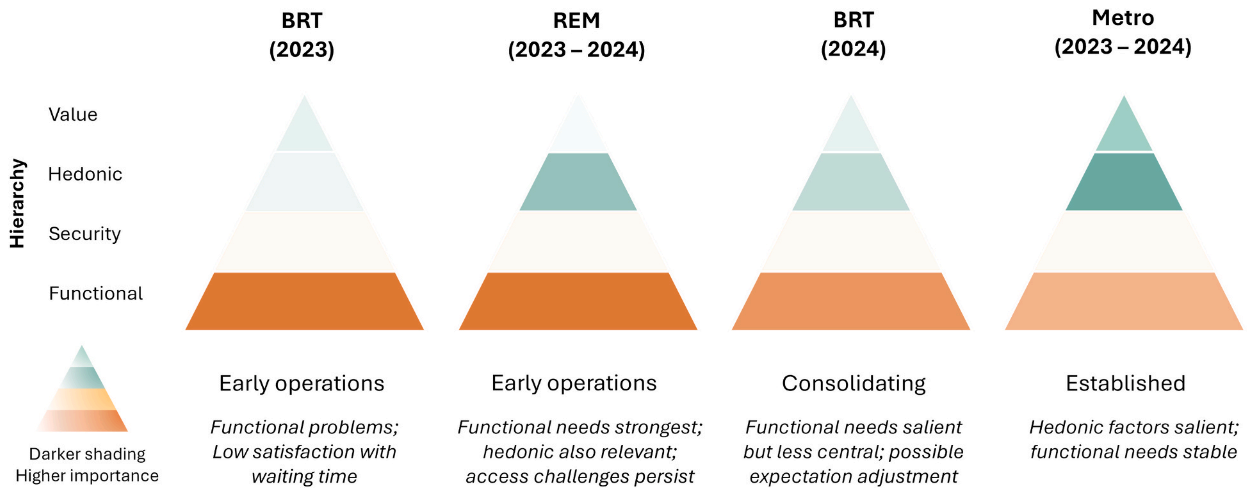

Fig. 7. Hierarchy of transit needs across modes and stages of operational maturity.

framework. While it provides a useful lens for interpreting satisfaction dynamics, its conditional relevance assumption does not always hold in practice. Because maturity varies primarily across systems in our case study, the patterns discussed here should be interpreted in terms of system-level differences rather than as within-mode evidence of maturity effects. Future research could explore how factors such as within-mode operational maturity, service familiarity, expectation adjustment, and user coping mechanisms contribute to shifting perceptions over time, particularly in newer systems.

7.3. Policy implications

Findings from this study suggest that, at the aggregate level, the drivers of user satisfaction evolve in tandem with the degree to which foundational needs are met. As such, planners and agencies should avoid focusing on higher-order features before ensuring that the basics, such as reliability, speed, accessibility, are consistently delivered (Diab et al., 2015). Improvements should be sequenced considering system-level operational maturity, with investments tailored to address salient needs.

In the case of Montreal's Metro, results indicate that satisfaction is increasingly shaped by comfort and experience-related features. For this mode, planners may consider prioritizing investments that enhance the user environment, including seating, cleanliness, ventilation, and crowd management. By contrast, the Pie-IX BRT (still undergoing infrastructure completion) remains constrained by unmet functional needs, particularly wait times. In this sense, improvements targeting operational consistency should take precedence before hedonic factors. Moreover, operational improvements for this route would benefit a low-income corridor, where riders are more likely to be transit dependent. In the REM, a new system experiencing frequent service disruptions, information quality emerged as a key driver of trip satisfaction. This points to the importance of real-time communication strategies, including onboard announcements, mobile notifications, and responsive customer service, to manage expectations.

While the salience of different satisfaction attributes varies by mode and by each system's operational maturity, some expectations appear universal. Travel time consistently emerged as a core determinant across all modes, underscoring that functional performance remains foundational to user satisfaction, even when other needs rise in prominence. Transit agencies should therefore continue to prioritize system designs and operational strategies that minimize travel times and improve perceived efficiency. In practice, strategies to improve travel time and efficiency can include reducing intersection delay through signal priority (particularly for BRT/LRT at-grade segments), shortening dwell times via all-door boarding and level boarding where applicable, and improving headway regularity through active operations control (e.g., holding, dynamic dispatching, and schedule adjustments). Because users experience is strongly influenced by service reliability and travel time variability (Soza-Parra et al., 2019; Weng et al., 2023), agencies should prioritize measures that reduce variability as much as average travel time.

Finally, the study underscores the importance of overall perceptions, particularly during the implementation phases of new infrastructure. In both REM and BRT systems, overall satisfaction with the system itself (beyond individual trip attributes) played a large role in shaping evaluations. This finding suggests that initial perception-building efforts, including design quality, branding, and communication campaigns, should be treated not as an afterthought, but as strategic investments. Strong first impressions can set the tone for how users assess performance in subsequent years, making early trust-building a core part of long-term satisfaction strategies.

8. Conclusion

This paper examined temporal changes in the drivers of trip satisfaction across Montreal's three rapid transit systems (i.e., the Metro, the Pie-IX BRT, and the REM light rail) over a one-year period. We focused on three central questions: (i) how these drivers evolved between 2023 and 2024, (ii) whether patterns differed across systems at different stages of operational maturity, and (iii) how the results align with the hierarchy of transit needs framework.

Our findings highlight differences in the drivers of trip satisfaction across Montreal's rapid transit systems. The metro exemplifies a long-established service in which satisfaction is stable and increasingly shaped by comfort and other higher-order attributes. By contrast, the BRT illustrates how newer systems can be subject to evolving expectations, with functional attributes such as waiting time initially dominating before gradually declining in salience as riders adapt. Meanwhile, the REM reflects an intermediate case, where both functional and hedonic considerations play parallel roles in shaping satisfaction.

With respect to theory, the results suggest only partial alignment with the hierarchy of transit needs. At the attribute level, inconsistencies appear, especially for newer systems where adaptation and shifting expectations complicate the assumed progression from functional to hedonic concerns. At the aggregate layer level, however, stronger support was found, with higher-order attributes gaining explanatory importance in the metro, while functional needs remained prominent for the REM and BRT. These findings indicate that the framework may serve best as a broad organizing principle rather than a strict rule, with its conditional relevance assumption proving more robust when applied at the categorical level than to individual factors.

This study is not without limitations. The analysis covers a relatively short period of one year, which may constrain the ability to detect longer-term stabilization or structural shifts in satisfaction drivers. In addition, the use of a four-point Likert scale forced respondents into non-neutral categories. Moreover, the number of trip attributes available and assessed is narrow, and the survey does not record which metro lines respondents used, and Montreal's metro lines are all long-established, limiting meaningful within-metro comparisons of maturity. Finally, we cannot directly test whether satisfaction drivers vary by riders' experience or habit strength with a given mode. While respondent random intercepts account for stable, time-invariant unobserved individual-level differences in baseline satisfaction, identifying whether more experienced riders weigh service attributes differently would require additional data not available in our dataset. Future research should address these limitations by extending the period of analysis, expanding the set of measured attributes, and examining whether the conditional relevance of different need layers depends more directly on the fulfillment of lower-order expectations. Such efforts will be essential to fully capture how satisfaction evolves as transit systems mature and adapt to changing travel behaviors.

CRediT authorship contribution statement

Thiago Carvalho: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ahmed El-Geneidy:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

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.

Acknowledgement

The authors would like to thank the anonymous reviewers for their feedback on the manuscript. The authors would like to thank Daniel Schwartz and Yasser Kazma from McGill IT for their help with the survey administration as well as the Transportation Research at McGill (TRAM) members Rodrigo Victoriano-Habit, Hisham Negm and Meredith Alousi-jones for their support with the data cleaning and preparation for analysis. This research was funded by Natural Sciences and Engineering Research Council of Canada grant Towards a better understanding of the determinants and satisfaction of travel among different groups in major Canadian Cities (NSERC RGPIN-2023-03852), Fonds de Recherche du Québec Société et Culture (https://doi.org/10.69777/370079), the Social Sciences and Humanities Research Council's partnership grant Mobilizing justice (SSHRC 895-2021-1009), the Canadian Institute of Health Research (CIHR PJT-195797), the Tier 1 Canada Research Chair in Public Transport Planning and Operation (CRC-2025-00098), and the Natural Sciences and Engineering Research Council of Canada (NSERC RGPIN-2023-03852).

Appendix

Appendix A1: Stepwise Bayesian multi-level ordered probit models of trip satisfaction.

Predictors	Base		Contextual		Transit Perceptions		Trip Attributes	
	Median	95% CrI	Median	95% CrI	Median	95% CrI	Median	95% CrI
Thresholds								
Dissatisfied | Satisfied	**−2.13**	**−2.54 – -1.76**	**−2.17**	**−3.02 – -1.38**	**−1.64**	**−2.39 – -0.94**	**−3.77**	**−4.68 – -3.08**
Satisfied | Very Satisfied	**1.15**	**0.83 – 1.47**	**1.43**	**0.66 – 2.27**	**2.02**	**1.32 – 2.77**	**2.26**	**1.77 – 2.84**
Mode *(Baseline: Metro)*								
BRT	**−0.50**	**−0.95 – −0.08**	**−0.87**	**−1.44 – −0.32**	**−0.68**	**−1.14 – −0.25**	−0.42	−1.00 – 0.13
REM	−0.12	−0.55 – 0.30	−0.34	−0.90 – 0.22	0.24	−0.18 – 0.67	−0.10	−0.66 – 0.45
Wave *(Baseline: 2023)*								
2024	−0.16	−0.47 – 0.15	−0.13	−0.45 – 0.19	−0.04	−0.37 – 0.28	−0.29	−0.73 – 0.13
BRT:2024	**0.45**	**−0.04 – 0.95**	**0.59**	**0.07 – 1.13**	**0.58**	**0.05 – 1.10**	**1.15**	**0.45 – 1.89**
REM:2024	0.19	−0.29 – 0.68	0.20	−0.29 – 0.71	−0.00	−0.52 – 0.50	−0.13	−0.82 – 0.55
Travel time (Baseline: > 60 min)								
< 15 min			**2.35**	**1.48 – 3.25**	**1.78**	**0.98 – 2.63**		
15 to 30 min			**1.16**	**0.64 – 1.71**	**0.83**	**0.35 – 1.33**		
30 to 45 min			**0.96**	**0.53 – 1.39**	**0.63**	**0.23 – 1.04**		
45 to 60 min			0.36	−0.05 – 0.77	0.23	−0.16 – 0.62		
Gender (Baseline: Women)								
Men			0.04	−0.46 – 0.54				
BRT: Men			0.32	−0.36 – 1.02				
REM: Men			**0.79**	**0.08 – 1.51**				
Age (Baseline: 65 and over)								
18 to 29			**−0.98**	**−1.73 – -0.24**	**−0.76**	**−1.41 – -0.13**		
30 to 44			**−0.96**	**−1.65 – -0.26**	**−0.70**	**−1.33 – -0.09**		
45 to 64			**−0.76**	**−1.46 – -0.04**	−0.43	−1.09 – 0.19		
Neighborhood changes								
Improvement in quality of life			**0.49**	**0.33 – 0.66**	**0.22**	**0.07 – 0.39**		
Public transit perceptions								
Quality of life					**0.35**	**0.17 – 0.53**		
Transit satisfaction					**0.40**	**0.23 – 0.58**	**0.26**	**0.07 – 0.47**
Transit improvement (past year)					**0.33**	**0.16 – 0.50**		
Trip attributes (Main mode)								

(*continued on next page*)

(*continued*)

Predictors	Base		Contextual		Transit Perceptions		Trip Attributes	
	Median	95% CrI	Median	95% CrI	Median	95% CrI	Median	95% CrI
Access time							**0.41**	**0.20 – 0.63**
Waiting time							**0.33**	**0.04 – 0.63**
BRT: Waiting time							**0.77**	**0.34 – 1.21**
REM: Waiting time							0.39	−0.07 – 0.86
Travel time							**0.82**	**0.57 – 1.11**
Comfort							**0.90**	**0.59 – 1.25**
BRT: Comfort							**−0.64**	**−1.08 – -0.21**
REM: Comfort							−0.08	−0.57 – 0.40
Information							**0.65**	**0.43 – 0.90**
Cost-benefit							**0.58**	**0.37 – 0.80**
Random Effects								
τ_{00}	0.12		0.19		0.25		0.31	
ICC	0.62		0.42		0.25		0.06	
N	354 $_{ID}$		354 $_{ID}$		354 $_{ID}$		354 $_{ID}$	
Observations	708		708		708		708	
Marginal R^2/Conditional R^2	0.006/0.469		0.162/0.521		0.312/0.535		0.652/0.724	

Note: Median reflects posterior medians (probit-scale coefficients, β); 95% CrI denotes the 95% credible interval.

Data availability

Data will be made available on request.

References

Agresti, A., Natarajan, R., 2001. Modeling clustered ordered categorical data: a survey. Int. Stat. Rev. 69 (3), 345–371.

Al-haj, E., Alamoush, S., Darwish, M., 2024. Service quality of Amman BRT based on user's perceptions: a structural equation modeling approach. KSCE J. Civ. Eng. 28 (8), 3506–3519.

Allen, J., Muñoz, J., Ortúzar, J., 2019a. On evasion behaviour in public transport: Dissatisfaction or contagion? Transportation Research Part A -Policy and Practice 130, 626–651.

Allen, J., Muñoz, J., Ortúzar, J., 2019b. Understanding public transport satisfaction: using Maslow's hierarchy of (transit) needs. Transp. Policy 81, 75–94.

Allen, J., Muñoz, J., Ortúzar, J., 2020. On the effect of operational service attributes on transit satisfaction. Transportation 47 (5), 2307–2336.

Anderson, E., Mittal, V., 2000. Strengthening the satisfaction-profit chain. J. Serv. Res. 3 (2), 107–120.

APTA. (2025). *Public transport ridership report: Fourth Quarter 2024*.

Arancibia, S., González, F., Busco, C., Vera, T., Yuretic, M., 2025. Recognizing user satisfaction and loyalty in bus and metro services: a gender-based analysis using PLS-SEM. Res. Transp. Bus. Manag. 59, 101322.

Baldwin, S.A., Fellingham, G.W., 2013. Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. Psychol. Methods 18 (2), 151–164.

Bürkner, P.-C., 2021. Bayesian Item Response Modeling in R with brms and Stan. J. Stat. Softw. 100 (5).

Calvo, E., Ferrer, M., 2018. Evaluating the quality of the service offered by a bus rapid transit system: the case of Transmetro BRT system in Barranquilla. Colombia. International Journal of Urban Sciences 22 (3), 392–413.

Cao, J., Cao, X., 2017. Comparing importance-performance analysis and three-factor theory in assessing rider satisfaction with transit. J. Transp. Land Use 10 (1), 837–854.

Cao, J., Cao, X., Zhang, C., Huang, X., 2016. The gaps in satisfaction with transit services among BRT, metro, and bus riders: evidence from Guangzhou. J. Transp. Land Use 9 (3), 97–109.

Carvalho, T., & El-Geneidy, A. (2024). Everything has changed: the impacts of the COVID-19 pandemic on the transit market in Montréal, Canada. *Transportation.*.

Carvalho, T., El-Geneidy, A., 2025. Measuring the impacts of a major metro disruption in Montréal, Canada, on riders' satisfaction and willingness to recommend the service to others. *Case Studies on.* Transp. Policy 101480.

Carvalho, T., Victoriano-Habit, R., MacNeil, H., Valero, A., & El-Geneidy, A. (2025). *Measuring the impacts of the Réseau express métropolitain (REM): Progress Report 2019-2024*.

News, C.B.C., 2024. Shorter REM trains running on weekends and holidays as of Sunday. CBC News. https://www.cbc.ca/news/canada/montreal/shorter-rem-trains-1.7178927.

News, C.B.C., 2025. Quebec's transport minister calling for solutions in wake of REM service disruptions. CBC News. https://www.cbc.ca/news/canada/montreal/rem-transportation-service-disruptions-1.7462099.

CDPQ Infra. (2017). *Réseau électrique métropolitain (REM) - Sommaire des prévisions d'achalandage du REM Février 2017*.

Chepuri, A., Arkatkar, S., Joshi, G., Pulugurtha, S., 2024. Assessing bus user's perception using structural equation modelling. Aust. J. Civ. Eng. 1–9.

Choi, S., Ko, J., Kim, D., 2021. Investigating commuters' satisfaction with public transit: a latent class modeling approach. Transp. Res. Part D: Transp. Environ. 99, 103015.

City News. (2023, 23/11/2023). Making Montreal's metro safer, new approach by police and STM. *City News Everywhere*. https://montreal.citynews.ca/2023/11/23/making-montreal-metro-safer-police-stm/.

Conlon, M., Foote, P., O'Malley, K., Stuart, D., 2001. Successful arterial street limited-stop express bus service in Chicago. Transp. Res. Rec. 1760 (1), 74–80.

Daykin, A., Moffatt, P., 2002. Analyzing ordered responses: a review of the ordered probit model. Underst. Stat. 1 (3), 157–166.

De Oña, J., de Oña, R., Eboli, L., Forciniti, C., Mazzulla, G., 2016. Transit passengers' behavioural intentions: the influence of service quality and customer satisfaction. Transportmetrica a: Transport Science 12 (5), 385–412.

De Oña, J., De Oña, R., Eboli, L., Forciniti, C., Mazzulla, G., 2018. An ordered regression model to predict transit passengers' behavioural intentions. *Case Studies on.* Transp. Policy.

De Oña, R., Machado, J., de Oña, J., 2015. Perceived service quality, customer satisfaction, and behavioral intentions: Structural equation model for the metro of Seville, Spain. Transp. Res. Record(2538), 76–85.

DeBell, M., & Krosnick, J. A. (2009). *Computing weights for American National Election Study survey data* (ANES Technical Report series, no. nes012427, Issue. http://www.electionstudies.org.

Deng, W.-J., Kuo, Y.-F., Chen, W.-C., 2008. Revised importance–performance analysis: three-factor theory and benchmarking. Serv. Ind. J. 28 (1), 37–51.
Diab, E., El-Geneidy, A., 2014. Transitory optimism: changes in passenger perception following bus service improvement over time. Transp. Res. Rec. 2415 (1), 97–106.
Diab, E.I., Badami, M., El-Geneidy, A., 2015. Bus transit service reliability and improvement strategies: Integrating the perspectives of passengers and transit agencies in North America. Transp. Rev. 35 (3), 292–328.
Diaz, A., Cantillo, V., Arellana, J., 2023. Understanding how individuals perceive changes in the built environment and the transport system after implementing a BRT system. the case of Barranquilla, Colombia. J. Transp. Geogr. 110, 103623.
Díez-Mesa, F., de Oña, R., de Oña, J., 2018. Bayesian networks and structural equation modelling to develop service quality models: Metro of Seville case study. Transportation Research Part A -Policy and Practice 118, 1–13.
Dillman, D., Smyth, J., Christian, L., 2014. *Internet, phone, mail, and mixed-mode surveys: the tailored design method*, (4th ed.). Wiley.
Dou, M., Gu, Y., Gong, J., 2024. How do people perceive the quality of urban transport service? New insights from online reviews of Shanghai metro system. Journal of Urban Management 13 (4), 705–719.
Efthymiou, D., Kaziales, M., Antoniou, C., Tyrinopoulos, Y., 2014. Measuring the effects of economic crisis on users' perceptions of public transport quality. Transp. Res. Rec. 2415 (1), 1–12.
Fielbaum, A., Tirachini, A., 2020. The sharing economy and the job market: the case of ride-hailing drivers in Chile. Transportation 48 (5), 2235–2261.
Fu, X., Juan, Z., 2017. Drivers of transit service loyalty considering heterogeneity between user segments. Transp. Plan. Technol. 40 (5), 611–623.
Fu, X., Zhang, J., Chan, F., 2018. Determinants of loyalty to public transit: a model integrating Satisfaction-loyalty Theory and Expectation-Confirmation Theory [Article]. Transp. Res. A Policy Pract. 113, 476–490.
Gärling, T., Axhausen, K., 2003. Introduction: Habitual travel choice. Transportation 30 (1), 1–11.
Gärling, T., Garvill, J., 1993. Psychological Explanations of Participation in Everyday Activities. Advances in Psychology.
Gelman, A., Carlin, J., Stern, H., Rubin, D., 2013. *Bayesian data analysis*. CRC Press.
Haider, M., Anwar, A., Mehbub, H., 2023. The prevalence of telework under Covid-19 in Canada. Inf. Technol. People 36 (1), 196–223.
Henriquez-Jara, B., Arriagada, J., Tirachini, A., 2025. Impact of real-time information on passenger satisfaction across varying public transport quality levels in 13 Chilean cities. Transp. Res. A Policy Pract. 200, 104622.
Huan, N., Hess, S., Yao, E., 2022. Understanding the effects of travel demand management on metro commuters' behavioural loyalty: a hybrid choice modelling approach. Transportation 49 (2), 343–372.
Ibrahim, S., Zhang, Y., Ahmad, A., 2025. Rail-based public transportation service quality and customer satisfaction: a decade of insights and advances. Public Transp. 1–25.
Johnson, V., Albert, J., 2000. *Ordinal data modeling*. Springer Science & Business Media.
Kim, S., Ulfarsson, G., 2012. Commitment to light rail transit patronage: Case study for St. Louis MetroLink [Article]. Journal of Urban Planning and Development-Asce 138 (3), 227–234.
Lai, W., Chen, C., 2011. Behavioral intentions of public transit passengers: the roles of service quality, perceived value, satisfaction and involvement. Transp. Policy 18 (2), 318–325.
Li, L., Gao, T., Yu, L., Zhang, Y., 2022. Applying an integrated approach to metro station satisfaction evaluation: a case study in Shanghai, China. Int. J. Transp. Sci. Technol. 11 (4), 780–789.
Liddell, T., Kruschke, J., 2018. Analyzing ordinal data with metric models: what could possibly go wrong? J. Exp. Soc. Psychol. 79, 328–348.
Luo, S., He, S., 2021. Using data mining to explore the spatial and temporal dynamics of perceptions of metro services in China: the case of Shenzhen. Environ. Plann. B: Urban Anal. City Sci. 48 (3), 449–466.
Machado, J., de Ona, R., Diez-Mesa, F., de Ona, J., 2018. Finding service quality improvement opportunities across different typologies of public transit customers. Transportmetrica A Transport Science 14 (9), 761–783.
Machado, L., Larranaga, A., Ruiz-Padillo, A., Nodari, C., Cybis, H., 2024. Quantifying the impact of metrorail transport quality factors on overall user satisfaction: a Brazilian case study. Transportation 1–44.
Makowski, D., Ben-Shachar, M., Lüdecke, D., 2019. bayestestR: describing effects and their uncertainty, existence and significance within the Bayesian framework. Journal of Open Source Software 4 (40), 1541.
Matzler, K., Bailom, F., Hinterhuber, H., Renzl, B., Pichler, J., 2004. The asymmetric relationship between attribute-level performance and overall customer satisfaction: a reconsideration of the importance–performance analysis. Ind. Mark. Manag. 33 (4), 271–277.
Mohammadi, M., Rahimi, E., Davatgari, A., Javadinasr, M., Mohammadian, A., Bhagat-Conway, M., Salon, D., Derrible, S., Pendyala, R., Khoeini, S., 2023. Examining the persistence of telecommuting after the COVID-19 pandemic. Transp. Lett. 15 (6), 608–621.
Nassereddine, M., Eskandari, H., 2017. An integrated MCDM approach to evaluate public transportation systems in Tehran. Transportation Research Part A -Policy and Practice 106, 427–439.
Oliver, R., 1976. Effect of expectation and disconfirmation on postexposure product evaluations: an alternative interpretation. J. Appl. Psychol.
Oliver, R., 1997. *Customer satisfaction: a behavioral perspective on the consumer*, (2nd Edition ed.). McGraw-Hill.
Osorio-Arjona, J., Horak, J., Svoboda, R., García-Ruíz, Y., 2021. Social media semantic perceptions on Madrid Metro system: using Twitter data to link complaints to space. Sustain. Cities Soc. 64, 102530.
Ouali, L., Graham, D., Barron, A., Trompet, M., 2020. Gender differences in the perception of safety in public transport. J. R. Stat. Soc. Ser. A Stat. Soc. 183 (3), 737–769.
Palm, M., Allen, J., Zhang, Y., Tiznado-Aitken, I., Batomen, B., Farber, S., Widener, M., 2022. Facing the future of transit ridership: shifting attitudes towards public transit and auto ownership among transit riders during COVID-19. Transportation 1–27.
Parasuraman, A., Zeithaml, V., Berry, L., 1985. A conceptual model of service quality and its implications for future research. J. Mark. 49 (4), 41–50.
Parasuraman, A., Zeithaml, V., Berry, L., 1988. SERVQUAL - a multiple-item scale for measuring consumer perceptions of service quality. J. Retail. 64 (1), 12–40.
Parker, M., Li, M., Bouzaghrane, M., Obeid, H., Hayes, D., Frick, K.T., Rodriguez, D., Sengupta, R., Walker, J., Chatman, D., 2021. Public transit use in the United States in the era of COVID-19: Transit riders' travel behavior in the COVID-19 impact and recovery period. Transp. Policy 111, 53–62.
Pasek, J., 2018. anesrake: ANES Raking Implementation. R Package Version 80. https://CRAN.R-project.org/package=anesrake.
Qi, Y., Liu, J., Tao, T., Zhao, Q., 2023. Impacts of COVID-19 on public transit ridership. Int. J. Transp. Sci. Technol. 12 (1), 34–45.
Rabe-Hesketh, S., Skrondal, A., 2008. *Multilevel and longitudinal modeling using Stata*. STATA press.
Rodriguez-Valencia, A., Ortiz-Ramirez, H., Simancas, W., Vallejo-Borda, J., 2022. Understanding transit user satisfaction with an integrated bus system. J. Public Transp. 24, 100037.
Saeidi, T., Mesbah, M., Habibian, M., 2020. Sequenced ordered logit model considering latent variables for determining trip satisfaction of metro passengers. Transp. Res. Rec. 2674 (9), 755–766.
Saleem, M., Afzal, H., Ahmad, F., Ismail, H., Nguyen, N., 2023. An exploration and importance-performance analysis of bus rapid transit systems' service quality attributes: evidence from an emerging economy. Transp. Policy 141, 1–13.
Saxena, A., Choudhury, B., Das Gupta, P., 2024. Travel satisfaction of bus rapid transit users in a developing country: the case of Bhopal City. India. Transportation Research Record 2678 (9), 869–885.
Shin, E., 2023. What can we learn from online reviews? Examining the reviews of Los Angeles metro rail stations. J. Plan. Educ. Res. 43 (2), 254–267.
Singh, H., Kathuria, A., 2023. Heterogeneity in passenger satisfaction of bus rapid transit system among age and gender groups: a PLS-SEM Multi-group analysis. Transp. Policy 141, 27–41.
Soltanpour, A., Mesbah, M., Habibian, M., 2020. Customer satisfaction in urban rail: a study on transferability of structural equation models. Public Transp. 12 (1), 123–146.

Soza-Parra, J., Raveau, S., Muñoz, J., Cats, O., 2019. The underlying effect of public transport reliability on users' satisfaction. Transportation Research Part A -Policy and Practice 126, 83–93.
Statistics Canada. (2023). *2021 Census of Population. Statistics Canada Catalogue no. 98-316-X2021001*. Retrieved May 24, 2023 from https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E.
Sun, S.C., Duan, Z.Y., 2019. Modeling passengers' loyalty to public transit in a two-dimensional framework: a case study in Xiamen, China [Article]. Transportation Research Part A -Policy and Practice 124, 295–309.
The Canadian Press. (2023, 01/11/2023). REM service disruptions Wednesday morning, third time in three days. *City News Everywhere*. https://montreal.citynews.ca/2023/11/01/rem-service-disruptions/.
van Lierop, D., Badami, M., El-Geneidy, A., 2018. What influences satisfaction and loyalty in public transport? a review of the literature. Transp. Rev. 38 (1), 52–72.
Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27 (5), 1413–1432.
Vehtari, A. G., J. , Magnusson, M., Yao, Y., Bürkner, P., Paananen, T., & Gelman, A. (2025). *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models*. In (Version R package version 2.9.0, https://mc-stan.org/loo/).
von Bergmann, J., Shkolnik, D., Jacobs, A., 2021. cancensus: R package to access, retrieve, and work with Canadian Census data and geography. R Package Version (4), 2. https://mountainmath.github.io/cancensus/.
Wan, D., Kamga, C., Hao, W., Sugiura, A., Beaton, E., 2016. Customer satisfaction with bus rapid transit: a study of New York City select bus service applying structural equation modeling. Public Transp. 8, 497–520.
Weng, J., Yu, J., Di, X., Lin, P., Wang, J., Mao, L., 2023. How does the state of bus operations influence passengers' service satisfaction? a method considering the differences in passenger preferences. Transp. Res. A Policy Pract. 174, 103734.
Yilmaz, V., Ari, E., Oğuz, Y., 2021. Measuring service quality of the light rail public transportation: a case study on Eskisehir in Turkey. Case Studies on Transport Policy 9 (2), 974–982.
Zhang, C., Cao, X., Nagpure, A., Agarwal, S., 2019. Exploring rider satisfaction with transit service in Indore, India: an application of the three-factor theory. Transp. Lett. 11 (8), 469–477.
Zhao, P., Gao, Y., 2022. Public transit travel choice in the post COVID-19 pandemic era: an application of the extended Theory of Planned behavior. Travel Behav. Soc. 28, 181–195.
Zheng, Y., Kong, H., Petzhold, G., Barcelos, M., Zegras, C., Zhao, J., 2021. User satisfaction and service quality improvement priority of bus rapid transit in Belo Horizonte. Brazil. Case Studies on Transport Policy 9 (4), 1900–1911.