Latent classes and attribute bounds in public transport route choice adaptations to unplanned metro disruptions

Lawrence Christopher Duncan *[1,2], Ouassim Manout [2]

[1] Department of Technology, Management and Economics, Technical University of Denmark, Copenhagen, Denmark

[2] LAET – ENTPE, CNRS, University of Lyon, Lyon, France

SHORT SUMMARY

Designing public transport systems resilient to unplanned disruptions requires understanding mode and route choice preferences during such events. This study, for the first time, uses Automatic Fare Collection data to analyse multimodal public transport route choice adaptations to unplanned metro disruptions. Methodologies are developed for identifying commuters and their usual commute route, approximating home/work locations, and identifying disrupted trips and onward post-disruption routes. These methodologies are operationalised in a case study of Lyon, France, where 4569 post-disruption routes are identified, including routes taking bridging buses (bus replacement services). A novel Latent Class Conjunctive Bounded Path Size (LC-CBPS) model is employed to model the route choices, inaugurally exploring both a) whether there are any latent classes of choice preferences, and b) whether travellers impose bounds upon different route attributes in their route consideration. Goodness-of-fit of the LC-CBPS model is benchmarked against other models, and estimated attribute preferences and bounds are analysed.

Keywords: automatic fare collection data, conjunctive bounded choice model, discrete choice modelling, disruption detection, public transport, route choice

1. INTRODUCTION

Designing public transport systems resilient to unplanned disruptions requires understanding mode and route choice preferences during such events. Several studies have explored preferences during disruptions through stated/revealed-preference surveys (e.g. Rahimi et al., 2019,2020; Drabicki et al., 2021; Bickel et al., 2026). These studies, however, tend to explore how respondents choose between public transport and other travel options, and do not assess preferences for different features of actual public transport routes.

Passive tracking data such as Automatic Fare Collection (AFC) data provides greater possibilities for analysing preferences for actual public transport route choices. The focus of most existing studies, however, has been on preferences during planned disruptions, e.g. Arriagada et al. (2025a,b), Mojica (2008), and Van den Berg (2025). Analysis of travellers' preferences during unplanned disruptions remains underexplored, despite their importance in designing public transport networks resilient to such events.

In the current study we, for the first time, use AFC data to analyse multimodal public transport route choices immediately after being disrupted by an unplanned metro service closure (post-metro-disruption), with an application to commuters in Lyon, France. Methodologies are

developed for identifying commuters and their usual commute route, approximating home/work locations, and identifying disrupted trips and onward post-disruption routes.

To model post-disruption route choices we employ a choice model that has not been applied before: a Latent Class Conjunctive Bounded Path Size (LC-CBPS) route choice model. The LC-CBPS model combines the well-known Latent Class (LC) choice model (McFadden & Train, 2000) with the recently developed Conjunctive Bounded Path Size (CBPS) model by Rasmussen et al. (2024) and Duncan et al. (2025). The LC model identifies multiple discrete latent classes of choice preferences, and the CBPS model estimates bounds that travellers impose on different route attributes, e.g. access/egress/transfer time. While latent classes of preferences have been explored in a handful of normal-condition public transport route choice studies (Arriagada et al., 2022,2025b; Liu & Duan, 2024; Bansal et al., 2022), they have not yet been explored in disrupted choices, other than in a recent stated preference survey (Bickel et al., 2026). The CBPS model has thus far only been applied to evaluate bounds on total route cost and local detouredness in car route choice (Duncan et al., 2025). It has yet to be applied to evaluate bounds on individual route attributes, which is very relevant public transport route choice, though so far unexplored.

The abstract is structured as follows. In Section 2 we develop a methodological framework for exploiting AFC data to model post-metro-disruption route choices. We also discuss the features of the LC-CBPS route choice model and how it is estimated/evaluated. In Section 3 we provide and discuss the estimation results, benchmarking the LC-CBPS model against other models and interrogating the estimated post-metro-disruption preferences. Section 4 concludes the abstract.

2. METHODOLOGY

Figure 1 illustrates the overall process of the methodology for modelling post-metro-disruption route choices.

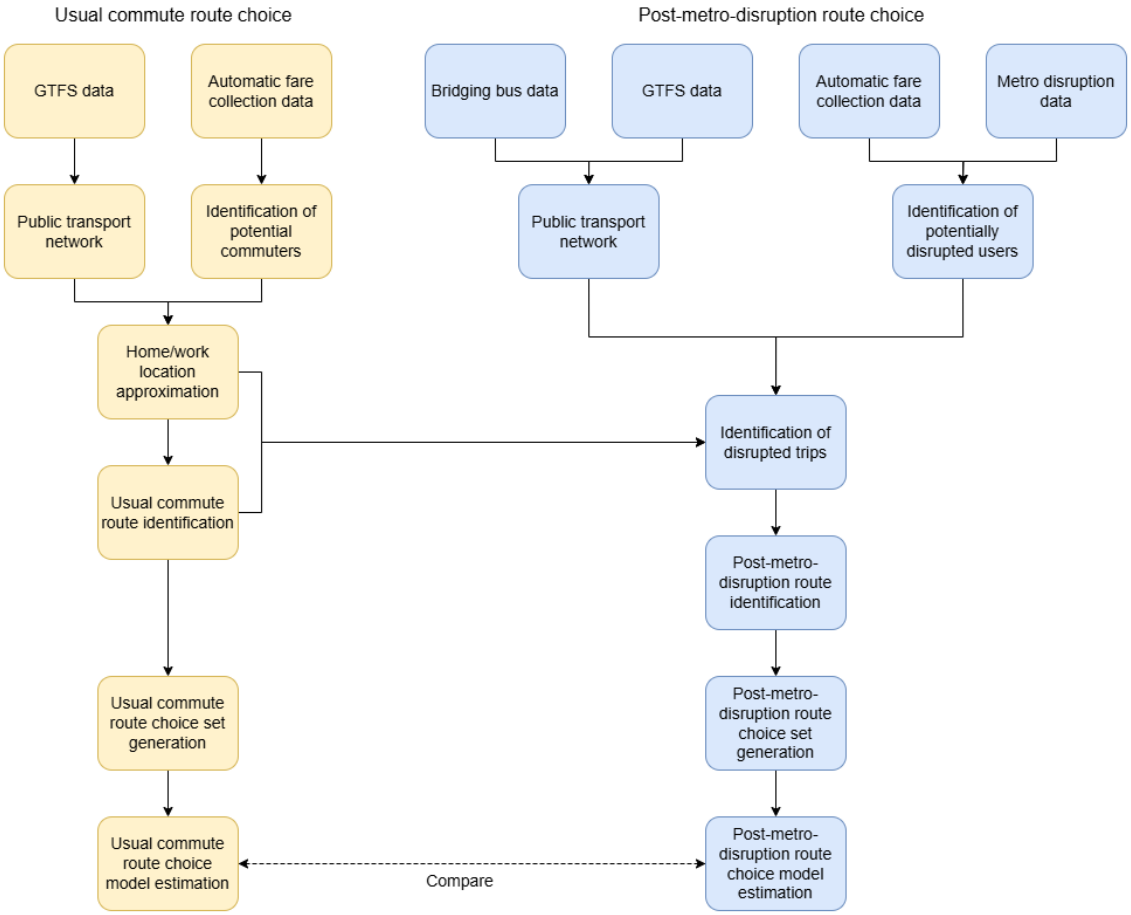


Figure 1. Schematic diagram of the overall process of the methodology.

2.1 Public transport network

Lyon's public transport network was obtained from General Transit Feed Specification (GTFS) data. A frequency-based model was set up using GTFS data from the 14th October 2022, the day within the study period 15th September - 15th October 2022 with the worst disruption and no missing services. Transfer links were added between all nodes (stops) within 600 meters haversine distance of each other, with a 4 km/h walking speed assumed. Waiting times were calculated as half the headway of the line and are hour-of-the-day dependent. There are four transport modes: bus, tram, metro, and bridging bus (bus replacement service). Bridging bus lines and headways were obtained from processing Automatic Vehicle Location (AVL) bridging bus data, and added to the network for the relevant day and hour-of-day. Figure 2 gives an example of a bridging bus line added to the network during a disruption.

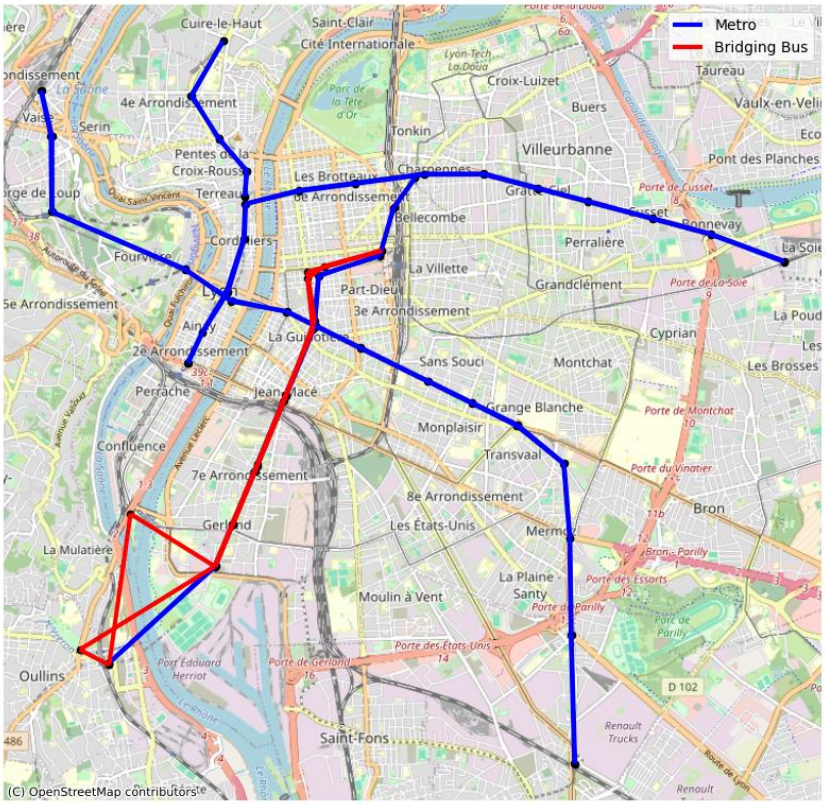


Figure 2. Metro network in Lyon and example of bridging bus line.

2.2 Identification of potential commuters

The study focuses on the travel behaviour of commuters, analysed through interrogating Lyon AFC data, which should be noted does not include check-out information. Stereotypical commuting behaviour was identified using a rule-based method. A day was flagged as a potential commute day for a user if:

- The first and last check-ins in the morning (6-10am) were no more than 2 hours apart.
- There were no daytime (10am-3pm) check-ins.
- The last check-in in the morning and the first check-in in the afternoon (3-8pm) were no less than 4 hours apart.

This identified potential commuters and potential commuting days.

2.3 Home/work location approximation

To more accurately identify the selection of first and last stops a commuter may consider, access/egress times, and disrupted trips, a method is proposed for approximating true home/work locations, based solely on exploiting AFC data. The first check-ins of a user in the morning are used to approximate home locations, and the first check-ins in the afternoon on flagged potential commute days are used to approximate work locations. We harness the stop locations of the first check-ins as well as the frequency each stop was used. The overarching assumption made is that travellers will walk maximally 600m haversine distance from their home to the first stop of their morning commute. The idea then is to find the location that lies 600m haversine distance between the first check-in stops. A mixed integer program was employed to identify the coordinate point

where the most frequently visited selection of stops can each be reached, and that is distanced between the selected stops and weighted closer to the most frequently visited stop(s). Figure 3 demonstrates the method. Figure 4 gives an example of the approximated home location from a user's check-ins.

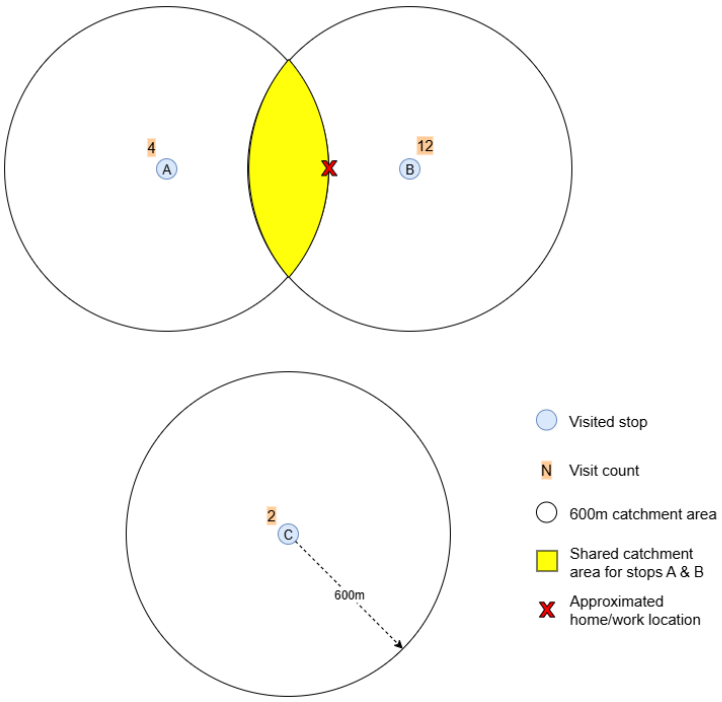


Figure 3. Demonstration of the method for approximating true home/work locations.



Figure 4. Example of approximated home location from user's check-ins.

2.4 Usual commute route identification

Having identified potential commuters and approximated their home and work locations, we then verified commuters and identified their usual commute routes. This was done by trip chaining the check-ins on flagged potential commute days, and then testing whether the trips corresponded with a feasible morning home-to-work commute and afternoon work-to-home commute return. Since check-out information is unavailable, the trip chaining was conducted through exploiting the GTFS network via conducting shortest path searches between check-ins, as well as to home/work location. The usual commute routes were obtained through concatenating the sub-routes generated from check-in to check-in and to destination. Figure 5 gives an example of usual home-to-work and work-to-home commute routes identified.

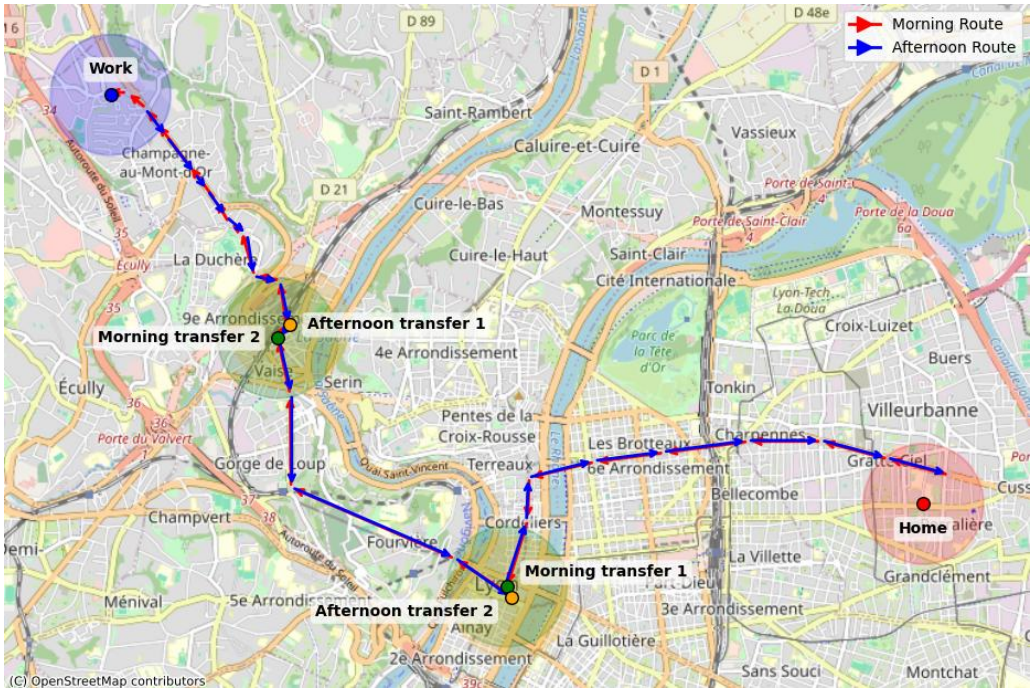


Figure 5. Example of usual home-to-work and work-to-home commute routes identified.

2.5 Identification of disrupted trips

Disruptions to the metro in Lyon during the study period were extracted from service disruption logs provided by the public transport operator KEOLIS Lyon. Only metro service closures that occurred on weekdays during commuting times and lasted longer than 20 minutes were considered. We then interrogated the AFC data to identify disrupted trips based on the service closure information. Disruption identification was categorised into four behaviour groups, based on check-in behaviour. This identified commuters that: i) entered the metro but realised it was closed, ii) travelled on the metro and experienced a disruption en route, iii) travelled on the metro aware of a disruption to a part of their usual route, and iv) travelled to their usual boarding metro stop but realised it was closed from outside. It was then verified that the check-ins post-disruption aligned with travelling home / to work. If a feasible post-disruption trip was identified, then its post-disruption route was generated. Figure 6 gives an example of an identified disrupted trip.

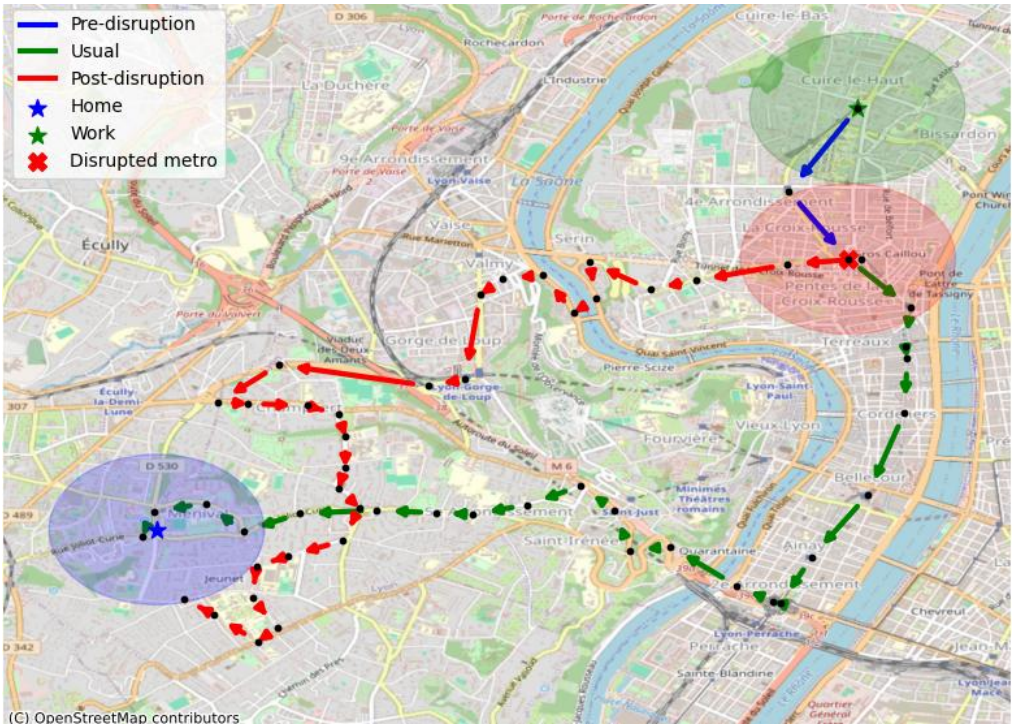


Figure 6. Example of an identified disrupted trip.

2.6 Route choice set generation

Post-metro-disruption route choice sets were generated by first randomly sampling pairs of boarding and alighting stops in the vicinity of the disrupted metro and home/work location, and then conducting shortest path searches based on different link weights. A combination of shortest path, labelling, simulation, and link elimination were employed to maximise the possibilities for generating routes with varying properties. Operating bridging bus lines were added to the network where appropriate and disrupted metro links were removed. Figure 7 displays an example of the route alternatives generated for a post-metro-disruption route observation. After removing some outliers and removing route observations where no generated route overlapped greater than 70% with the observed route, 4569 post-disruption route observations remained.

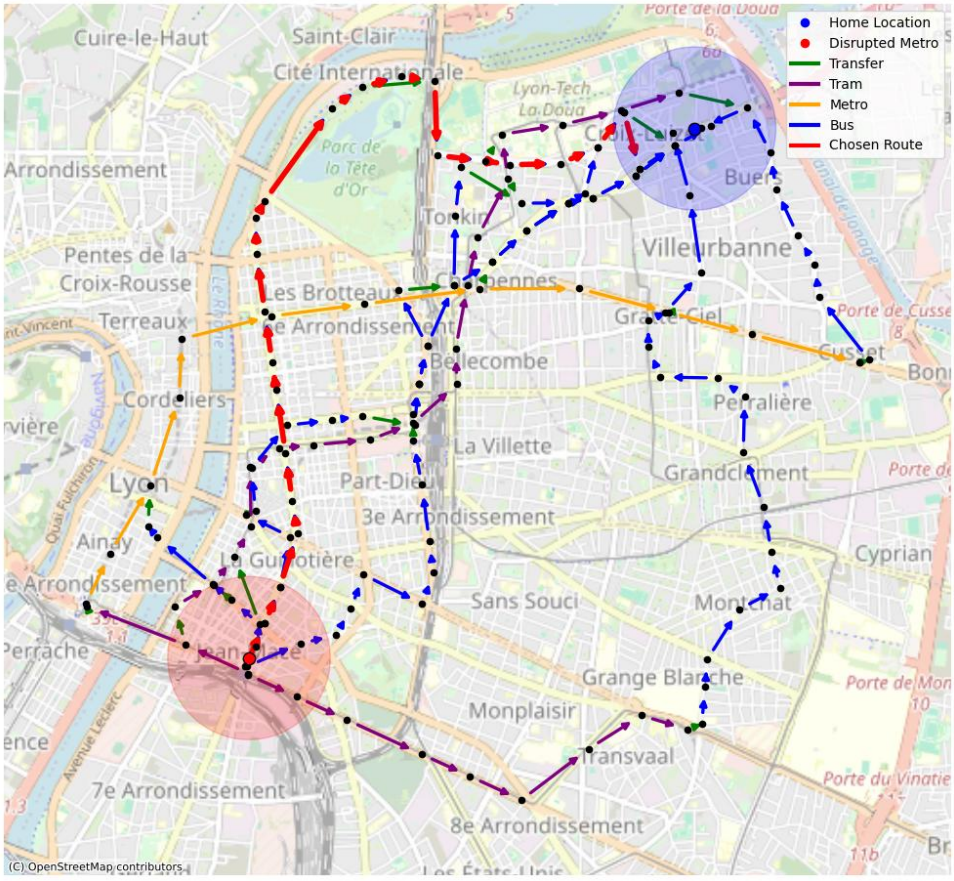


Figure 7. Example of the route alternatives generated for a post-metro-disruption route observation.

2.7 Route choice model

For the route choice model, we employ a LC-CBPS model, which adopts a CBPS model for each latent class. The CBPS model is a combination and adaption of the Conjunctive Bounded Choice Model (CBCM) derived from first principles in Section 3 in Rasmussen et al. (2024) and the Bounded Path Size Local Detour Threshold model formulated in Section 3 of Duncan et al. (2025). Word constraints in this extended abstract inhibit us from formulating the LC-CBPS model mathematically, which is not quick to do. We therefore summarise the features of the model as follows:

- The overall choice probability of a route is equal to the probability the individual belongs to a class, multiplied by the route choice probability if an individual belongs to that class. In this case, the former are estimated as parameters, and the latter is a CBPS probability function.
- Route attributes that are supposed to be potentially bounded follow a Bounded Choice Model (Watling et al., 2018) contribution to the CBPS probability function. As such, if a route has an attribute value greater than an attribute bound parameter away from the best (lowest) attribute value across routes in the choice set, then it receives zero choice probability. A route only needs to violate one attribute bound to be defined as non-considered (zero probability).

- Route attributes that are not supposed to be bounded follow a Multinomial Logit (MNL) contribution to the CBPS probability function.
- A bounded path size correction term is included within the CBPS probability function that penalises routes for overlapping with non-violating routes, i.e. capturing correlations between just the routes defined as considered.

The post-metro-disruption route attributes considered in this study are:

1. Number of transfers
2. In-vehicle time [min]
3. Access walking time [min] (from disrupted metro to boarding stop)
4. Access waiting time [min]
5. Transfer walking time
6. Transfer waiting time
7. Egress time (form alighting stop to approximated home/work location)
8. Tram mode-specific constant (MSC)
9. Metro MSC
10. Bridging bus MSC
11. Reconnection with usual commute (binary indicator)
12. Route robustness (measure of whether multiple lines service stops of route)

The attributes supposed to be potentially bounded are attributes 1-7, excluding access walking time as no potential for bounding was found empirically.

In this study, we suppose there are two latent classes of choice behaviour. For each of these classes, the CBPS model requires estimating the following parameters:

- A scaling parameter for each attribute.
- An absolute surplus attribute bound parameter for each bounded attribute.
- A bounded path size scaling parameter.

An additional parameter is estimated for the probability an individual belongs to Class 1. Parameters are estimated using a modified maximum likelihood estimation procedure, see Section 5 in Duncan et al. (2025). Parameter estimate significance is evaluated through conducting Likelihood Ratio Tests. The scaling parameters for the bounded attributes are tested against -0.001 (an arbitrarily small negative value), against infinity for the bound parameters, and zero otherwise. Attribute preferences are evaluated through computing Aggregate Semi-Elasticities (ASEs), which calculate the average percentage change in probability per unit attribute change.

3. RESULTS AND DISCUSSION

First, Table 1 benchmarks the goodness-of-fit of the LC-CBPS model against other relevant models. As can be seen, the LC-CBPS model provides by far the best fit to the data, indicating that: i) there is heterogeneity in preferences/bounds among users, captured well with two classes, ii) it is more realistic to capture correlations between just routes considered by users, and iii) attribute bounds exist and applying them to the generated routes improves behavioural realism.

Model	MNL	LC-MNL	CBCM	LC-CBCM	CBPS	LC-CBPS
Log-Likelihood	-6294	-5950	-6061	-5937	-5720	-5530
BIC	12698	12127	12282	12203	11600	11389
Adjusted rho squared	0.576	0.598	0.591	0.605	0.614	0.626
Number of parameters	13	27	19	39	19	39

Table 1. Goodness-of-fit metrics from estimations of the different models.

Table 2 displays results from estimating the LC-CBPS model, in terms of parameter estimates, attribute ASEs, and percentage of generated routes cut off by each bound. The first thing that stands out is that many of the attribute scaling parameters are statistically insignificant. This may appear problematic, but these are all for the bounded attributes where attribute scaling is also influenced by the attribute bound parameter. Considering this, all attributes were statistically significant for each class.

Number of transfers. Class 2 has a larger in magnitude number of transfers ASE than Class 1 (-0.968 compared to -0.750), suggesting that Class 2 are strongly averse to transferring after a disruption.

Travel time. The ASEs indicate that, apart from for access walking time, Class 2 have weaker dispreferences for the different travel time attributes. This aligns with Class 2 being very transfer averse and caring less about travel time. Interestingly, the transfer walking and waiting time ASEs for Class 2 (-0.076 and -0.072) are less negative than the ASEs for Class 1 (-0.142 and -0.135). This suggests Class 1's strong aversion to transferring is independent from the travel time involved in transferring. This could be because these travellers perceive transferring as coming with a greater risk of further disruption, and are particularly fearful of it after already experiencing a disruption. Class 2 having a stronger dispreference for access walking time (ASE of -0.082 compared to -0.042 for Class 1) could be linked to Class 2 having a lower dispreference for bridging buses over bus (ASE of -0.335 compared to -0.886 for Class 1). Bridging buses are generally located close to the disrupted metro stops, and thus access walking averse travellers are likely to select bridging bus.

Mode. The ASEs indicate that tram is preferred over bus by both classes, metro is preferred over bus by Class 2 but not Class 1, and neither class prefer bridging bus over regular bus.

Reconnection. The reconnection ASE is positive for Class 1 and negative for Class 2, implying that Class 1 have a preference for reconnecting with their usual commute after a disruption, while Class 2 have a dispreference. This aligns with Class 1 having stronger preferences for bus and tram and Class 2 having stronger preferences for metro and bridging bus. An exploration found that reconnecting with usual commuting route was linked with taking bus and tram, but not metro or bridging bus.

Robustness. Both classes have a positive preference for robustness, suggesting that after disruptions travellers prefer to take routes that are more robust to further disruption. The robustness ASE for Class 1 (0.901) is greater than the ASE for Class 2 (0.371), implying that Class 1 have a stronger preference for route robustness. This aligns with Class 1 having stronger preferences for bus and tram, as these are the modes that can have multiple options between stops.

Attribute bounds. Both classes are not willing to take more than 1 surplus transfer after a disruption. For the other attributes, there are some minor differences in the estimated bounds between classes. Most notably, Class 1 do not impose any bound on surplus egress time, while Class 2 impose a bound of 8.2 minutes. Interestingly, the bounds for surplus transfer waiting time (6-7 minutes) are tighter than the bounds for surplus access waiting time (9-10 minutes). This suggests that after being disrupted travellers are more lenient in the threshold they impose upon waiting to access than waiting to transfer. The most restrictive surplus travel time bound is on transfer walking time: travellers are not willing to walk more that 5.6-5.9 minutes more than the best option to transfer after a disruption. Overall, the different estimated bounds on surplus route attributes defines 50.3% of the generated routes as unconsidered, which is a large proportion.

4. CONCLUSION

This extended abstract has, for the first time, exploited AFC data to explore public transport route choice adaptations to unplanned disruptions. Methodologies were developed for identifying and modelling disrupted route choices, and an advanced choice model was employed to model route choice preferences, novelly exploring both latent classes of preferences and bounds imposed on attributes. Although not provided here due to space constraints, we have also compared the estimated attribute preferences/bounds with other choice models and with non-disrupted commute route choices. These will be presented at the conference. Overall, the results have the potential to be harnessed for a range of system/policy design purposes, such as evaluating network vulnerability, designing networks, or optimising bridging bus services.

ACKNOWLEDGEMENTS

We gratefully acknowledge the funding provided by the European Union under the Horizon Europe program for the project “PATH2ZERO” [Grant ID: 101184132]. We are also thankful to KEOLIS for providing the data on public transport use in Lyon, and to Sytral Mobilité for providing historical GTFS data.

Parameter	Estimate		ASE		Percentage of Routes Cut		
	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2	Total
Number of transfers	-1.240	-3.494*	-0.750	-0.968			
In-vehicle time	-0.036	-0.001*	-0.063	-0.031			
Access walking time	-0.059	-0.197	-0.042	-0.082			
Access waiting time	-0.001*	-0.001*	-0.109	-0.055			
Transfer walking time	-0.001*	-0.001*	-0.142	-0.076			
Transfer waiting time	-0.001*	-0.001*	-0.135	-0.072			
Egress time	-0.327	-0.461	-0.239	-0.215			
Tram MSC	0.754	1.769	0.558	0.481			
Metro MSC	-2.703	0.806	-0.924	0.283			
Bridging bus MSC	-2.318	-0.780	-0.886	-0.335			
Reconnection	1.532	-1.537	1.049	-0.587			
Robustness	1.264	0.892	0.901	0.371			
Bounded Path Size	0.811	1.882					
Bound (number of transfers)	1.996	1.005			28.6%	28.6%	28.6%
Bound (in-vehicle time)	25.01	24.32			3.9%	4.3%	3.9%
Bound (access waiting time)	9.055	10.10			4.5%	1.2%	1.2%
Bound (transfer walking time)	5.693	5.835			18.1%	17.5%	17.5%
Bound (transfer waiting time)	6.971	6.351			22.5%	26.5%	22.5%
Bound (egress time)	13.71*	8.156			0%	2.5%	0%
Class membership probability	0.360	0.640					
Total routes cut					51.6%	53.3%	50.3%

Table 2. Estimation results for the LC-CBPS model. * = not significant at 95% level.

REFERENCES

Arriagada J, Munizaga M, Guevara C, & Prato C, (2022). Unveiling route choice strategy heterogeneity from smart card data in a large-scale public transport network. Transportation research part C: emerging technologies, 134, p.103467.

Arriagada J, Guevara C, Munizaga M, & Gao S, (2025a). An experiential learning-based transit route choice model using large-scale smart-card data. Transportation, 52(4), p.1543-1568.

Arriagada J, Prato C, & Munizaga M, (2025b). Incorporating the inertia effect into a route choice model using fare transaction data from a large-scale public transport network. Transportation Research Part A: Policy and Practice, 196, p.104467.

Bansal P, Hörcher D, & Graham D, (2022). A dynamic choice model to estimate the user cost of crowding with large-scale transit data. Journal of the Royal Statistical Society Series A: Statistics in Society, 185(2), p.615-639.

Bickel J, Geržinič N, Van Oort N, De Bruyn M, & Molin E, (2026). Heterogeneity in route choice behaviour during unplanned train disruptions considering the possibility of teleworking. Transportation Research Part A: Policy and Practice, 203, p.104708.

Drabicki A, Islam M, & Szarata A, (2021). Investigating the impact of public transport service disruptions upon passenger travel behaviour—Results from Krakow city. Energies, 14(16), p.4889.

Duncan L, Rasmussen T, Watling D, Cazor L, & Nielsen O, (2025). Identification of local detouredness and correlation effects in a bounded route choice model: solution & estimation in large-scale networks. Transportmetrica A: Transport Science, p.1-49.

Liu X & Duan Z, (2025). Exploring factors affecting metro route choice behavior: a hybrid latent class modeling approach. Transportation Planning and Technology, 48(8), p.1635-1653.

McFadden D & Train K, (2000). Mixed MNL models for discrete response. Journal of applied Econometrics, 15(5), p.447-470.

Mojica C, (2008). Examining changes in transit passenger travel behavior through a smart card activity analysis (Doctoral dissertation, Massachusetts Institute of Technology).

Rahimi E, Shamshiripour A, Shabanpour R, Mohammadian A, & Auld J, (2019). Analysis of transit users' waiting tolerance in response to unplanned service disruptions. Transportation Research Part D: Transport and Environment, 77, p.639-653.

Rahimi E, Shamshiripour A, Shabanpour R, Mohammadian A, & Auld J, (2020). Analysis of transit users' response behavior in case of unplanned service disruptions. Transportation Research Record, 2674(3), p.258-271.

Rasmussen T, Duncan L, Watling D, & Nielsen O, (2024). Local detouredness: A new phenomenon for modelling route choice and traffic assignment. Transportation Research Part B: Methodological, 190, p.103052.

Van den Berg K, (2025). Analyzing route choice preferences in response to planned public transport disruptions.

Watling D, Rasmussen T, Prato C, & Nielsen O, (2018). Stochastic user equilibrium with a bounded choice model. Transportation Research Part B: Methodological, 114, p.254-280.