Investigating the Long-Term Impact of Disruptions on Passenger Travel Behavior Using AFC Data

A Case Study of Washington D.C. Metro Network

by

Fariba Tavakoli

in partial fulfillment of the requirements for the degree of

Master of Science

in Transport, Infrastructure and Logistics

at the Delft University of Technology,

to be defended publicly on Monday August 19, 2024 at 10:30 AM.

Student number:	5716632	
Thesis committee:	Prof. Dr. O. (Oded) Cats	TU-Delft, Committee Chair
	Dr. Ir. Y. (Yongqiu) Zhu	TU-Delft, Supervisor
	Dr. Ir. M. (Maarten) Kroesen	TU-Delft, Supervisor

TUDelft

Preface

This thesis marks the end of my two-year adventure in the Transport, Infrastructure, and Logistics program at TU Delft. It's been quite the journey, and I'm incredibly grateful to everyone who helped me along the way.

First off, a huge thank you to my daily supervisor, Yongqiu. Your constant support, dedication, and brilliant ideas were a guiding light throughout this whole process. I feel really lucky to have had you as my supervisor. Thank you Oded, for all the honest feedback and for sparking my interest in public transport networks during your lectures. Your influence played a big role in shaping this research. Maarten, I'm so grateful for your critical suggestions and for helping me figure out the structure and methodology of this thesis. You were always there to answer my questions, and I really appreciate that. I couldn't have asked for a better committee, and I've learned so much from all of you during this journey.

To my dear family, I couldn't have done this without your support and motivation. To my amazing cheerleader, Ali, thank you for being by my side and supporting me through all the challenges of the past six months. And to my amazing friends, your motivation and support in this journey meant a lot to me.

I hope the readers of this thesis find it interesting and insightful.

Fariba Tavakoli
Den Haag, August 2024

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Definition
AFC	Automated Fare Collection
AVL	Automatic Vehicle Location
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
GMM	Gaussian Mixture Model
LCA	Latent Class Analysis
LMM	Latent Markov Model
MLMM	Mixture Latent Markov Model
MNL	Multinomial Logit
PT	Public Transport
RP	Revealed Preference
SP	Stated Preference

1 Introduction

Urban public transit systems play a crucial role in providing accessibility and mobility for society. They facilitate the daily commutes of millions of people and are essential for ensuring that individuals can reach their destinations efficiently. However, disruptions such as delays and cancellations are increasingly common in these systems due to various factors, including system failures, natural disasters, the growing need for maintenance on aging infrastructure, and construction work (Zhu et al., 2017). These disruptions can be either planned or unplanned. Unplanned disruptions can be caused by incidents such as signal failures, operational problems, accidents, weather, and vandalism (Liu et al., 2021; Shires et al., 2019). Planned disruptions typically arise from pre-scheduled activities such as maintenance closures, strikes, and similar events (Rahimi et al., 2020).

These disruptions can negatively affect urban mobility, leaving travelers who rely on public transport stranded and causing considerable inconvenience. Vulnerable groups, including the elderly and disabled, might face increased difficulties in accessing their destinations under such conditions. This can result in a loss of confidence in the public transport network. Disruptions can also lead to economic and opportunity losses, as well as increased vulnerability within the transit network. For instance, a rail traffic control malfunction at Amsterdam Central Station in the Netherlands on June 4, 2023, caused a major blockage that lasted several days, restricting passengers' access from Amsterdam to Schiphol Airport Station, one of Europe's most critical airports by passenger volume (Schiphol, 2022).

Due to the frequent disruptions, passengers might shift to alternative modes of travel, such as cars. If these alternatives are found more convenient, passengers may continue using them even after the disruptions are resolved (Karlaftis et al., 2006; Zhu et al., 2017). Cars have always been competitors to transit systems due to their higher comfort, privacy, availability, and shorter travel times. The emergence of ride-hailing and car-sharing services has increased the threat of alternative modes attracting more transit users. A shift towards car ridership in urban areas could worsen environmental and congestion issues as more people opt for cars.

1.1. Research Gap and Contributions

The existing literature extensively explores the immediate effects of disruptions on passenger travel behavior, evaluating the influence of passenger-related, journey-related, disruption-related, and network-related factors on passenger responses. However, research into the longer-term effects, extending at least five to six months post-disruption, is less prevalent and predominantly relies on qualitative methods such as interviews and surveys, which are limited in sample size, time-consuming, and expensive.

Automatic fare collection (AFC) systems offer a significant opportunity to analyze passenger behavior comprehensively. These systems record detailed journey information for all passengers, overcoming sample size limitations and reducing biases. Among the studies utilizing smart card data, most examine the impact during or immediately after the disruption, with only a few extending their analysis to two or three months post-disruption. Consequently, there is a gap in investigating changes in passenger behavior over a longer period after the disruption using AFC data. One of the challenges of looking into the long-term is that passenger behavior might not remain uniform throughout the post-disruption period thus their behavior could vary within that relatively long period.

This study aims to address this gap by extending the post-disruption period to five or six months and dividing it into several sub-periods. This approach allows for a more granular analysis of passenger behavior, enabling us to capture and understand the evolving patterns in the post-disruption period using AFC data. Understanding the long-term impacts is important because it reveals the extent to which temporary changes in passenger behavior due to disruptions become permanent over time. Additionally, it allows for an examination of passengers' tolerance towards disruptions and the extent to which they maintain their travel behavior, demonstrating travel behavior inertia despite having experienced disruptions. If results indicate that passengers leave the system or reduce their usage after a disruption, the findings can guide the need for implementing mitigating measures not only immediately after disruptions but also in the following months to prevent users from abandoning the public transport (PT) in favor of other modes. This can include providing additional discounts to passengers significantly affected by major disruptions, especially those living in the suburbs who rely on PT.

Scientifically, this study contributes to the existing literature by developing a general framework which can be applied to other case studies to discover the long-term effects of disruptions on travel behavior. The method used in this study is a mixture latent Markov model, a novel approach for studying the long-term impact of disruptions. This model is specifically ideal in this study because it is suitable for clustering longitudinal data. In addition, it probabilistically assigns the data to clusters in contrast to deterministic clustering methods such as K-Means, thereby accounting for uncertainties.Moreover, this model is capable of uncovering unobserved travel patterns among passengers and their specific likelihood of altering their travel behavior over time. This powerful feature of the model is instrumental in identifying behavior changes induced by disruptions.

1.2. Research Objective and Questions

The research objective is to gain a deep comprehension of the prolonged effects of public transport disruptions on passengers travel behavior. To attain this objective, the main research question is formulated as:

Main Research Question: How can we identify passengers' travel behavior change in the long-term due to disruptions?

This main question is answered by the following sub-questions:

1. What are the indicators that capture travel behavior regarding travel frequency, travel time regularity, time-of-day, and day-of-week?
2. How can we identify the differences among passengers regarding their travel behavior?
3. How can we measure the changes in travel behavior of passengers post-disruption compared to the pre-disruption period regarding travel frequency, travel time regularity, time-of-day, and day-of-week?

1.3. Thesis Organization

The structure of this document is organized as follows: Chapter 2, the literature review, provides an overview of the existing research and identifies current gaps. Chapter 3 details the methodology used to answer the research questions, and Chapter 4 is dedicated to the explanation of the case study and the data sets which are used in the analysis. Chapter 5 presents the analysis results from the case study. Finally, Chapter 6 delves into a discussion of the results and provides insights for future research and recommendations.

2 Literature Review

2.1. Search Strategy

The sources used to find relevant papers for the literature review were Scopus and Google Scholar. The main keyword combinations that we used were ”passenger/traveller”, ”travel behavior/pattern”, ”disruption” together with either of ”public transport”, ”public transit”, ”rail”, and ”metro”. Majority of these paper resulted in studies that were focused on the immediate impact, therefore, later we also added the term ”long term impact” to find the papers with a focus on longer term impacts. We excluded papers that considered COVID-19 as the disruption because our study primarily focuses on public transport disruptions. Additionally, papers whose main focus was building an estimation model for predicting passenger choices rather than exploring behavior and influencing factors were also excluded. Moreover, several additional papers were identified using the backward snowballing technique.

This search resulted in 18 papers. The main focus, analysis method, data source, case study, mode, and main findings with regards to travel behavior from these papers are summarized in Table 2.1.

Following the line of research in Eltved et al. (2021) and Nazem et al. (2018) who evaluated passenger behavior using segmentation and clustering, we decided to enrich our understanding of papers that used clustering as the main approach to evaluate passenger behavior. Therefore, we employed keyword combinations such as ”passenger segmentation/clustering” or ”travel behavior segmentation/-clustering” and ”AFC or smart card data.” The synthesis of the papers on travel behavior segmentation is found in Section 2.3.

Table 2.1: Overview of the reviewed papers.

Study	Main Focus	Analysis	Data Source	Case Study	Mode	Main Finding w.r.t Travel Behavior
Mo et al. (2022)	Network performance, passenger flow, and mode choice under unplanned delays	Used graph theory to analyze resilience, vulnerability, and redundancy, used binary logit model to analyze travel mode choices	AFC, AVL	Chicago, US	Rail	Passengers switched to alternative lines in high redundancy areas, and to bus in low redundancy areas. Reduced fare status passengers were more likely to stick with PT. High-income passengers were more likely to use alternative modes like ridesharing services.
Tian and Zheng (2018)	Commuter behavior during unexpected train delays (UTD)	Used DBSCAN to extract commuters' regular travel patterns, and compared the trips affected by a disruption to these regular travel patterns	AFC	Singapore	Rail	86% continue with the disrupted mode (by waiting or postponing their departure time), and 14% switch to other modes (bus, taxi, or bike). Key factors influencing commuter responses included the cause of UTD, time of day (peak vs. non-peak), and whether the UTD occurred en-route or pre-trip.
Liu et al. (2021)	Evaluate network performance and passengers' behavior under unplanned disruptions	Compared system performance metrics (e.g., travel times, passenger volumes) between typical days and disruption days, infer passenger responses such as travel continuation, delayed departures, altered OD pairs, and mode changes (e.g., shuttle bus usage).	AFC, AVL	Not disclosed	Metro	Staying within the disrupted mode was more favorable than switching modes, and most passengers maintained their usual travel schedules and ODs. The passengers outside the system when disruptions happen are more likely to change their origin/destination station, their mode, and departure time compared to those inside. The passengers whose origin and destination stations are affected are most likely to change their mode.
Zhu et al. (2017)	Passenger responses during planned disruptions (capacity reduction and station closure)	Analyzed the differences between stated preferences and actual choices, and the impact of socio-demographics	SP[1]-RP[2] surveys	Washington D.C, US	Metro	Wealthier riders more likely to switch modes compared to low-income groups who preferred sticking to regular bus services, 20% of respondents did not return even after full service restoration.
Rahimi et al. (2020)	Passenger decisions during an unplanned disruption and the influencing factors	Random parameter multinomial logit (RPMNL) model was used to account for heterogeneity and panel effects. Socio-demographic, personal attitude, trip-related, and built-environment variables tested for significance.	SP-RP surveys	Chicago, US	Metro, rail, and bus	Using own vehicle is more favorable for those whose destination is considered long-distance. Waiting time of the replacement bus negatively affects its utility. Those in suburban areas, prefer waiting for the replacement bus due to limited alternatives and their higher cost. Those with higher education and younger people are more likely to choose ride-sharing.
Murray-Tuite et al. (2014)	Passengers' behavior changes due to a deadly metro collision and the relationships between these changes and individual/travel characteristics	Employed multinomial logit (MNL) models to estimate probabilities of making specific behavioral changes.	Surveys	Washington DC, US	Metro	Not making any changes had the highest utility, followed by avoiding the front and rear rail cars, and then changing modes. Cost savings and frequent delays after the incident increase the likelihood of changing modes. Frequency of Metrorail use before the collision did not significantly affect mode change likelihood, but frequent users were less likely to change seats.
Eltved et al. (2021)	Changes in travel behavior before, during, and after a long-term rail line closure	Used K-means to group and compare passengers travel behavior before and after the disruption	AFC	Copenhagen, Denmark	Rail	Eight distinct travel behavior clusters were identified, after the disruption 12% of frequent commuters of the affected line did not return and a significant amount decreased their travel frequency.
Saxena et al. (2019)	Passenger behavior under delays vs. cancellations, and the difference in how they weigh wait time, travel time, and cost differently	Used a latent class choice model with socio-demographic and trip-specific attributes to capture variations in preferences under delays vs. cancellations	SP-RP surveys	Chicago, US	Rail	Travelers perceive cancellations three times more inconvenient than delays. Travelers are more likely to continue using PT in case of delays, and are less likely to use replacement buses in case of cancellation.
Drabicki et al. (2021)	Travel behavior shifts and travel information usage during unplanned disruptions	Used statistical tests and exploratory analysis to analyze the behavior.	SP-RP surveys	Krakow, Poland	Bus and tram	Majority (39%) changed their route, 29% waited at the stop, and the rest either walked or shifted to private modes. Passengers made long term adjustments (change routes or departure times) due to frequent PT disruptions. Frequent PT users were less likely to change their travel routine and more likely to wait at the stop.

Study	Main Focus	Analysis	Data Source	Case Study	Mode	Main Finding w.r.t Travel Behavior
Papangelis et al. (2016)	The effects of public transport disruptions on passenger behavior in rural areas	Used grounded theory and inductive content analysis, focus groups with bus/car users and cyclists, and interviews with bus drivers.	Semi-structured interviews and focus groups	The Scottish Borders, The United Kingdom	Bus	Passengers made permanent changes (avoiding certain travel times or routes) in response to frequent disruptions and lifestyle changes (buy cars/relocate) in response to long-term disruptions. Short-term disruptions led to minor adaptations (change modes/routes).
Li, Yao, Yamamoto, Huan, and Liu (2020)	Passenger choices under unplanned metro disruptions and the determinant factors	Used nested logit model (upper level for mode shift decisions (shift or non-shift) and lower level for specific travel plan choices)	SP survey	Guangzhou, China	Metro	A longer disruption duration, disruption occurring in peak, and unavailability of alternative routes increases the likelihood of shifting to other modes. Females are more likely to change modes. People with higher income shift modes more often as they have a higher value of time.
Nguyen-Phuoc et al. (2018b)	Mode choice, and factors influencing the travel behavior of passengers during a complete cancellation of all PT modes for an entire day	Applied MNL and nested logit models to analyze mode choice and behavior during PT disruptions. Examined factors such as gender, income, trip purpose, travel cost, transfer inconvenience, car ownership, driver's license ownership, and station accessibility	SP-RP surveys	Melbourne, Australia	Train, bus, tram	52% of users switch to driving a car, 31% to cycling or walking or cancel their trips, 11% to being a car passenger, and 5% to using taxi/Uber. Car/driver's license ownership and trip related trip purpose strongly influenced the preference for car use.
Nguyen-Phuoc et al. (2018a)	The factors influencing the choice of private car in the event of short-term and long-term disruptions rather than PT	Transcribed interviews and analyzed them using grounded theory principles	Semi-structured interviews	Melbourne, Australia	Train, tram	In a hypothetical ten-year disruption people stated willingness to buy a car, relocate, and find a job near their home. In this scenario, the key factors are travel time, distance, cost, destination, and flexibility. In a one-day disruption people consider changing their route, mode, or cancel their trip. In addition to the previous factors, car/driving license ownership, income, and trip purpose were significant in short-term disruptions.
Arslan Asim et al. (2021)	Mode choice behavior during short-term planned service disruptions (PSDs), and the influencing factors	Estimated a mixed MNL model and assessed the influence of travel time, wait time, and cost on the likelihood of choosing different modes	SP survey	Calgary, Canada	Light rail transit (LRT)	Ridership dropped by 35% during the disruptions. Transit customers with payment plans (students, older people, and low income) and weekend users more often continue using LRT during PSDs. Decreased wait time for express shuttles increased its appeal for passengers.
Shires et al. (2019)	How planned maintenance disruptions influence passengers' decisions to continue their journey using the disrupted mode, alternative modes/routes, reschedule or cancel their trips	Estimated MNL to examine choices and the way the different factors affect them	SP-RP surveys	London, The United Kingdom	Rail	While passengers are relatively adaptable, the type of replacement service, the management of connections and disruption periods significantly influence their decisions. Replacement buses result in higher demand loss compared to rail diversions. Awareness of disruptions does not significantly alter passenger behavior.
Adelé et al. (2019)	The behavior of suburban train passengers in response to disruptions of varying causes and severity	Used ascending hierarchical clustering (AHC) to define homogeneous groups of users with similar behavioral patterns during disruptions.	RP survey	Paris, France	Rail	Most passengers changed their arrival/departure stations or used other modes (car, bus, metro, tram) during major disruptions (delays of at least 30 minutes) while only a few opted to wait. Passengers with time constraints were less willing to wait. Those who received disruption information before starting their journey were more likely to change modes or routes.
Nazem et al. (2018)	The changes in ridership of disrupted stations and individual travel patterns due to a planned medium-term station closure	Daily ridership after the disruption was compared to the normal level. Travel pattern of frequent passengers was analyzed based on their used mode sequences (metro and bus)	AFC	Montreal, Canada	Metro	Ridership dropped at the closed stations during the disruption and increased at adjacent stations. Post-disruption, ridership at the closed stations did not fully recover to pre-closure levels. Frequent users showed a reduction in overall PT usage during and after the closures.

Study	Main Focus	Analysis	Data Source	Case Study	Mode	Main Finding w.r.t Travel Behavior
Li, Yao, Yamamoto, Tang, and Liu (2020)	Understand metro passengers' choice behavior under unplanned service disruptions, and how socio-demographic and travel characteristic attributes influence these choices	Error component latent class model (LCM) to capture the heterogeneity in travel plan choices under conditions of uncertainty.	SP survey	Guangzhou, China	Metro	Two latent classes are found: (1) uncertainty pessimists (high-income, female, and older) who are sensitive to disruption duration. (2) Uncertainty optimists (lower-income, male, younger) who are less sensitive to the disruption duration. Frequent travelers are more likely to belong to class 2.
This study	Understand how passengers' behavior changes in the long-term after disruptions	Estimating a mixture latent Markov model (MLMM) to cluster passenger travel behavior before and after a disruption (divided into several periods) based on a number of indicators.	AFC	Washington DC, US	Metro	Five travel patterns and three mobility styles are found. Unlike expectations, differences between affected and reference passengers were minimal. The affected group show a travel behavior inertia similar to the reference. The largest pattern, frequent less flexible peak travelers, exhibit the highest inertia (85%) among the affected passengers. Tendency towards decreased travel frequency was not significant among the affected.

[1]SP: Stated preference
[2]RP: Revealed preference

2.2. Passenger Responses to Disruption

The majority of the literature on the impacts of transit disruptions on passenger behavior focus on the immediate impacts, with only a limited number of studies exploring the longer term consequences. For example, Nazem et al. (2018) studied the effect of a planned station closure of three months on the ridership and passenger travel frequency. The results showed that even four months after the disruption, the level of ridership of the disrupted station did not return to the level under the normal situation. The ridership of adjacent stations which had increased during the closure did not revert back either; thus, suggesting long-term impacts on station ridership. Moreover, travel frequency of the frequent passengers decreased during the closure and did not increase to its pre-disruption level even two months after the closure was resolved. However, it should be investigated whether the findings of the paper would be valid for longer periods or not.

Additionally, Eltved et al. (2021) investigated the impacts of a planned line closure that lasted for three months. The study found that the level of ridership on the affected line did not return to its normal level even three months after the disruption had resolved. The author suggests that this is an indication of a long-term effect. Moreover, they found that among passengers who are considered to be regular commuters (people who travel everyday during peak-hour from Monday to Friday) 23% never returned to the system after the service resumed. The study speculates that when these commuters, who account for more than half the trips per day, face a disrupted service and switch to alternative modes to commute, they might find the new mode a better alternative to the public transport, and consequently not return to the system after the disruption.

Drabicki et al. (2021) uses a SP-RP surveys to study the short-term and long-term passenger behavior under unplanned disruptions and finds that 77% reported that they made long-term changes in their travel behavior as a result of experiencing frequent disruptions. These adjustments included changing their usual bus or tram routes to avoid routes that are frequently disrupted or changing their departure time. Additionally, 20% reported increased car usage.

Shires et al. (2019) uses a RP survey to understand passengers response to the disruptions caused by planned engineering works in the past 12 months. Majority of the passengers (64%) stated that they made no difference to their journeys due to the disruptions, and 18% stated that they are less likely to travel via the places they experienced disruptions. The limitations of the study is that the majority of the planned engineering works takes place in the weekend so the majority of the respondents were non-commuters. Additionally, the RP survey did not ask about the magnitude of the disruptions that passengers had experienced thus the author is not sure if there were any major disruptions to lead to long term impacts. Moreover, most of these people had experienced only one disruption in the past 12 months which can impact the outcome.

Among the few studies on the long-term effects, it is also worth mentioning the study by Murray-Tuite et al. (2014) which examines the effect of the Washington D.C. metro fatal accident in 2009 on travellers' long-term mode choice and the factors influencing their behavior. To capture the long-term effect, a survey was conducted in three waves spanning over a one-year period after the accident. The results of the study revealed that 10% of the respondents changed their travel mode and 17% changed their seating location on the train following the collision.

2.2.1. Factors Influencing Passenger Responses to Disruptions

The responses that passengers exhibit in the event of a disruption, can vary between changing the mode, the route, the departure time/station, waiting for the service to resume, and cancelling the journey. These responses are influenced by the context and the characteristics of the individual, the journey, the network, and the disruption itself. This section synthesizes the findings of the papers which are mentioned in Table 2.1 in more detail and analyzes the passenger responses in relation to the influencing factors.

Travel Purpose

Several studies have found that when the purpose of a journey is critical and urgent (e.g., work, school), passengers are more likely to change their routes (Adelé et al., 2019; Drabicki et al., 2021) or switch to modes other than PT, such as cars and taxis (Nguyen-Phuoc et al., 2018b; Mo et al., 2022; Papangelis et al., 2016; Tian and Zheng, 2018; Nguyen-Phuoc et al., 2018a). Additionally, they are less likely to cancel their journeys (Rahimi et al., 2020), which is likely due to the mandatory nature of the trip, time inflexibility, and the higher urgency to arrive on time.

Travel Duration

For trips of longer durations and distances, passengers are more likely to change their modes (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018) and switch to driving (if they own a car) (Rahimi et al., 2020; Nguyen-Phuoc et al., 2018b; Nguyen-Phuoc et al., 2018a). This behavior might be because the uncertainty of arrival time increases when PT is disrupted; therefore, passengers switch to other modes, like cars, which are more convenient and faster.

Disruption Time (peak, off-peak)

Disruptions occurring during morning and evening peak hours increase the likelihood of passengers shifting to other modes compared to off-peak disruptions. This may be because passengers need to arrive at their destinations, such as work or school, on time or because they are tired and need to arrive home earlier (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018). Additionally, Li, Yao, Yamamoto, Tang, and Liu (2020) found that if a disruption happens during peak hours, passengers are more willing to spend money to reduce travel costs, including travel time, disruption duration uncertainty, minimum disruption duration, and transfer time.

Income and Employment

People with higher incomes are more likely to change their mode of transportation (Li, Yao, Yamamoto, Huan, and Liu, 2020; Rahimi et al., 2020), switch to driving their own cars, or use car-sharing alternatives like Uber and Lyft (Zhu et al., 2017; Saxena et al., 2019), or take taxis (Li, Yao, Yamamoto, Tang, & Liu, 2020). Low-income groups, students, and the elderly generally continue using the PT or use the back-up shuttle bus, which is likely because urban transit is more affordable for them (Arslan Asim et al., 2021; Zhu et al., 2017; Mo et al., 2022).

Travel Frequency

Passengers who use PT more frequently are more likely to continue using the transit system in the event of a disruption, due to their high level of familiarity with the network, which enables them to navigate and find alternative routes more easily (Papangelis et al., 2016; Mo et al., 2022). Passengers with high travel frequency demonstrate a lower sensitivity to disruption uncertainties and exhibit more adaptable behavior towards unplanned service disruptions (Li, Yao, Yamamoto, Tang, & Liu, 2020).

Availability of Alternative Routes

In areas where alternative, undisrupted routes within PT modes are available, the majority of the demand can be accommodated by these alternatives. For example, Rahimi et al. (2020) found that in suburban areas, passengers opt for the backup shuttle bus when trains are disrupted. However, the lack of alternative routes increases the likelihood of passengers opting out of PT and choosing other modes (Li, Yao, Yamamoto, Huan, and Liu, 2020; Mo et al., 2022). In suburban areas with a limited supply of public transport, there is a lower likelihood that passengers change their routes in response to a disruption (Adelé et al., 2019).

Disruption Duration

Using a semi-structured survey, Nguyen-Phuoc et al. (2018a) found that when the disruption is short-term (at least one day), people consider changing their route and mode or canceling their journey. However, longer disruption durations increase the likelihood of shifting to other modes (Li, Yao, Yamamoto, Huan, & Liu, 2020).

Disruption Frequency

Drabicki et al. (2021) found that among passengers who frequently experience PT disruptions 77% stated that they made long-term adjustments to their behavior such as changing their route or departure time to avoid the routes/times that are likely to be disrupted. Frequent disruptions likely push passengers to find more reliable travel patterns. Similarly, in a study by Papangelis et al. (2016) passengers stated that due to frequent disruptions they made significant and permanent behavioral changes, such as mode changes, relocation, and job changes in the past.

Waiting Time of the Alternative Mode

The long waiting time for replacement buses can decrease the probability of choosing this mode and encourage passengers to turn to alternatives other than transit (Arslan Asim et al., 2021; Rahimi et al., 2020). Furthermore, additional inconveniences associated with these buses, such as walking from the platform and overcrowding, increase the likelihood that passengers switch to other modes (Zhu et al., 2017). Bus replacements were found by Shires et al. (2019) to be a less preferred mode in the case of planned disruptions. The results show that the loss in rail demand is three times higher when replacement buses are provided compared to rail diversions.

Driver’s License and Car Ownership

Passengers with a driver’s license or access to a car are more likely to change their mode and switch to driving (Nguyen-Phuoc et al., 2018b; Adelé et al., 2019). These individuals also have the option to rent a vehicle, whereas others, especially students who are less likely to own a driver’s license, would continue using public transport or cancel their trip in response to a disruption (Nguyen-Phuoc et al., 2018a).

Gender

Male passengers are more likely than the female passengers to wait for the service. Females are more sensitive to the uncertainty caused by the disruption and are more willing to pay for the reduction of travel time, minimize disruption duration, uncertain disruption duration, and transfer times (Li, Yao, Yamamoto, Tang, & Liu, 2020).

2.2.2. Methods in Current Literature

The methods used in the papers which are reviewed in this study are either survey-based or smart card data based. In order to investigate the travel behavior, mode choice, trip characteristics, and socio-demographics, the most traditional method to gather data is using surveys. However, there are a number of limitations to the RP and SP surveys. For example, the SP surveys can be unreliable because what passengers state they would choose under a hypothetical situation might differ from what they actually choose in reality. Moreover, the sample size is limited and the collection of the data requires a lot of investment and time (Liu et al., 2021).

A more recent method that can be seen in some of the papers is data-driven analysis using automated fare collection (AFC) data from passengers’ smart cards. Smart card readers inside stations or vehicles record a transaction each time a passenger taps their card, capturing the time, location, and card ID. While some systems require passengers to tap in and tap out, others only require tapping in. Unlike surveys, use of AFC data would require the inference of individual choices, i.e., whether passenger waited in the system, changed their routes, etc. Moreover, unlike surveys, socio-demographic information is not available in the AFC transaction and needs to be provided separately by the public transport authorities if available.

2.3. Passenger Segmentation

To study long-term changes in behavior due to disruptions, segmentation is employed. This method allows us to test our expectations about how the behavior of different passenger profiles might change. For instance, we might expect frequent travelers to reduce their travel frequency after a disruption, or passengers to avoid traveling at times when they previously experienced disruptions. Segmentation helps identify passenger profiles, enabling an efficient analysis of passenger behavior without the need to investigate each individual’s behavior in detail. This approach also avoids the shortcomings of aggregate-level evaluations, which can miss nuanced differences.

By forming behavioral clusters that capture passenger behaviors across different periods, we can track the transition of passengers among these clusters over time. This provides a better understanding of how the behavior changes.

To segment passengers, a set of indicators should be defined to capture the varying aspects of travel behavior. The choice of indicators is guided by our expectations about behavior changes, data availability, and existing literature. In this section, we provide an overview of some papers that use clustering methods to group passengers and analyze their behavior, synthesizing these papers based on the indicators they use for travel behavior.

2.4. Travel Behavior Indicators

One of the most frequently used indicators of travel behavior in the literature is the share of days with at least one journey made by a passenger, often referred to as ”active days.” This indicator has been widely employed in various studies as a measure of travel intensity and regularity. For instance, studies by Eltved et al. (2021), Wang et al. (2023), and Ma et al. (2013) have utilized active days, among other indicators, to segment passengers and create user profiles using K-Means clustering based on smart card data. Additionally, Ou, Cai, et al. (2018) applied an affinity propagation algorithm to segment passengers using several indicators, including active days. Ma et al. (2017) also used this indicator to analyze spatiotemporal regularities and patterns in commuting behavior through the iterative self-organizing data analysis technique (ISODATA).

In addition to active days, Eltved et al. (2021) also used the share of weeks with at least one journey made by passengers to further capture travel intensity and regularity.

Several studies have utilized the boarding time of journeys as an indicator of travel regularity. For instance, Ma et al. (2013) calculates the number of weekdays on which a passenger’s first journey of the day occurred at a similar time to characterize travel behavior. Expanding on this, Ma et al. (2017) also considers the consistency in the boarding times of passengers’ last journeys of the day. Similarly, Ghaemi et al. (2017) employs boarding time as a key metric for segmenting passengers using a generative model-based clustering approach.

Moreover, both Bhaskar, Chung, et al. (2014) and Medina (2018) use DBSCAN to identify regular boarding times for each passenger, aiming to uncover habitual time patterns. Medina (2018) further integrates the duration of activities (such as work or study) with boarding time to infer travel purposes. Additionally, Briand et al. (2017) uses a Gaussian mixture generative model to explore passengers’ temporal activity patterns based on disaggregated time series of their boarding times and the day of the week.

Another approach is presented by Zhao et al. (2017), who captures temporal travel patterns by dividing the day into three-hour slots and counting the number of days in which journeys were made during the same time slot for each passenger. Furthermore, Lathia et al. (2013) clusters passengers based on the similarity of their boarding times across weekdays. This study divides the day into five time bins and calculates the number of journeys a passenger made in each bin, grouping passengers with similar frequency vectors accordingly.

Additionally, several studies define temporal features to differentiate between journeys on weekdays and weekends or exclude weekends from the analysis altogether. For instance, Wang et al. (2023) incorporates the number of active weekdays and the number of journeys on weekdays as key features in the study, while Eltved et al. (2021) includes the share of journeys made on weekends alongside other indicators.

In addition to these temporal features, other studies have utilized various metrics such as the number of journeys with at least one transfer (Ou, Cai, et al., 2018), total number of journeys, and average travel time (Wang et al., 2023). Finally, while some studies also include spatial features, such as similar boarding stations, similar alighting stations, and similar route sequences (Bhaskar, Chung, et al., 2014; Ma et al., 2013, 2017; Wang et al., 2023), these aspects are not the focus of this study and are therefore not discussed in this literature review.

2.5. Conclusion and Discussion

By reviewing the literature, we find that a substantial body of work exists on the immediate impact of transit disruptions on passenger travel behavior. These immediate responses include changing modes, routes, departure times and/or stations, waiting, or canceling journeys. The factors influencing these behaviors are diverse: trip-related factors such as travel purpose, duration, and frequency; disruption-related factors such as the time, duration, and frequency of the disruption; service-related factors like the availability of alternative routes and waiting times for alternative modes; and passenger-related factors such as income, employment, car ownership, and gender. The literature review helps establish what behavioral changes can be expected among passengers as a result of a disruption and provides the foundation for identifying relevant travel behavior indicators.

Despite extensive research on the immediate impacts, there is a gap in understanding the long-term effects, defined as at least five to six months post-disruption. Studying these long-term impacts reveals whether the temporary changes in passenger behavior due to the disruption persist in the long run and to what extent passenger behavior changes or remains stable over time. In this study, we use AFC data to gather information on passenger journeys, and we divide the post-disruption period into several subperiods to obtain a detailed view of the evolution of behavior. We employ clustering to identify travel patterns based on various indicators across these periods. By examining passengers' membership in these behavior groups and their transitions over time, we aim to propose a framework to study the long-term impact of a disruption.

3 Methodology

In this chapter, the methodology for addressing the sub-questions is outlined which can also be seen in Figure 3.1. The first step is to identify the disruption instance and filter the passengers who are affected by it to study their behavior. In order to account for external factors such as seasonal trends, which might affect behavior regardless of the disruption, a group of passengers is considered as reference and the behavior of the affected passengers is compared to that of the reference. Next, we compare the behavior of each passenger before the disruption to after the disruption. We divide the post-disruption period into several periods for a more granular view of the behavior change. To understand behavior in each period, we calculate a set of indicators, e.g., average number of journeys, share of active days and etc, based on the smart card transactions of each passenger in each period. The result of this is a row of data which summarizes a passenger's behavior in a period. Next, these passenger behaviors are clustered using a mixture latent Markov model. With this method, a number of clusters are formed each showing a travel pattern. We can trace the membership of each passenger in these travel patterns across the periods, calculate the probability of shifting from one travel pattern to another over time, test our expectations against the actual outcome, and therefore conclude about passengers' behavior changes.

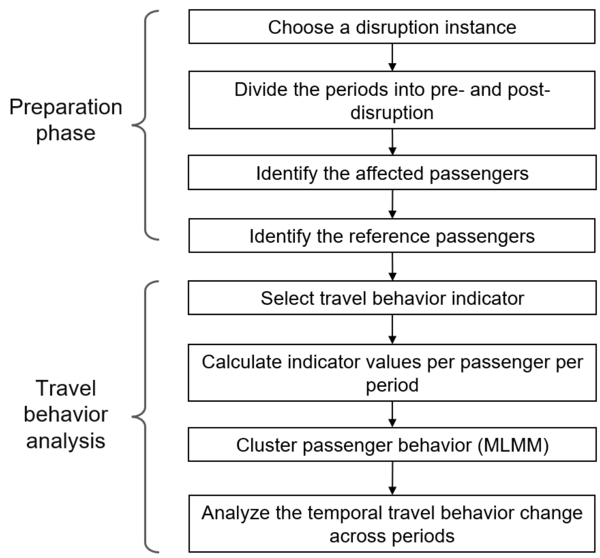

Figure 3.1: Methodology Steps

This chapter is organized as follows: The process of identifying the disruption instance and the division of the analysis periods are discussed in Sections 3.1 and 3.2, respectively. Next, the identification of the affected passengers and the reference passengers is explained in Sections 3.3, and 3.4. Section 3.5 provides an explanation about the mixture latent Markov model (MLMM) and 3.6 details the application of MLMM to segment passengers based on a set of temporal features which are extracted from Automated Fare Collection (AFC) transactions. Following that, the examination of passenger behavior change due to disruptions is elaborated upon in Section 3.7.

3.1. Disruption Identification Process

In this section, we explain the criteria for identifying a suitable disruption for the analysis. Subsequently, we discuss the process of exploring three datasets—the planned disruptions log, the unplanned disruptions log, and the AVL dataset—with the objective of finding a disruption. A suitable disruption is one which is most likely to have a lasting impact. For example, a single disruption of one hour would probably not have such impact. Therefore, the chosen disruption should be one which lasts for at least a few days or weeks. Alternatively, it can consist of multiple disruptions of smaller scale (a few hours or less) which happen repeatedly in a rather short period.

3.1.1. Disruption Identification Criteria

Several factors must be considered when identifying a suitable disruption. These factors lead to specific criteria that restrict the process of selecting the optimal disruption event. Each factor, along with the corresponding criterion it leads to, is outlined below:

The primary focus of this study is on passengers who predominantly travel during weekdays and working hours, as this group is more likely to consist of commuters rather than tourists who use the network sporadically. Therefore, the first criterion is:

- The disruption should occur on weekdays (excluding public holidays) and before 9 PM.

It's important that the disruption has the potential for a lasting impact on passenger behavior. This leads to the second criterion:

- The disruption should ideally last several days/weeks in a station or a set of adjacent stations. Multiple repeated disruptions of a few hours or less are also suitable.

Considering the significant decline in public transport usage during the Covid-19 pandemic (WMATA, 2024a), this study focuses exclusively on the pre-pandemic period, leading to the third criterion:

- The disruption must occur during the pre-Covid period, spanning from 1 August 2019 (when the data becomes available) to 1 March 2020.

To effectively assess the impact of the disruption, it's necessary to compare passenger behavior before and after the event, which informs the fourth criterion:

- There must be at least three weeks prior to the disruption, as the pre-disruption period, and a minimum of four or five months following the disruption for the post-disruption analysis.

Finally, to ensure the disruption's impact is isolated, meaning no other events skew the data, the final criterion is:

- There should be no other significant disruptions in the affected area during both the pre- and post-disruption periods.

Considering the second and third criteria, the effective period for identifying a disruption is between 21 August and 1 October 2019. This timeframe allows for sufficient pre- and post-disruption analysis without overlap from the Covid pandemic.

3.1.2. Finding a Disruption from the Planned/Unplanned Disruptions Log

First, we analyze the planned disruptions log to identify major maintenance works within the network that result in extended disruptions, either through reduced service frequency or complete closure of a line or station. These disruptions must meet the previously established criteria. Additionally, we consult the official WMATA website and review news coverage of the metro system to ensure that all significant planned maintenance works have been accounted for. Furthermore, we examine the unplanned disruptions log as an alternative source of disruptions.

It is important to recognize that disruptions may impact stations in one direction while the opposite direction remains unaffected. Therefore, during our analysis, we must carefully distinguish between disruptions occurring in opposite directions at the same station.

3.1.3. Finding a Disruption from the AVL Data

As a third source for identifying the occurrence of disruptions, we analyze the AVL dataset. One method to detect disruptions based on train movements involves calculating the time interval between the departures of two consecutive trains at a station. A significant interval suggests a delay at that particular station. To assess the occurrence of delays, we compute two specific measures based on the departure times of trains:

- **Headway per line per direction:** This metric calculates the time difference between the departure of two consecutive trains that visit the same station and travel in the same direction on the same line.
- **Headway per direction:** This metric measures the time difference between the departure of two consecutive trains at a station traveling in the same direction, regardless of their line.

Whenever large values are observed for both of these two measures at the same time, we can conclude that a significant delay has occurred.

The rationale for incorporating the second measure is that some stations are served by multiple lines, and while one line may be disrupted, others might operate normally. Therefore, depending on their destination, some passengers could remain unaffected by the disruption due to the availability of alternative lines. Consequently, observing a large value for the first measure while the second remains small does not necessarily signify a disruption for all passengers, as the availability of alternative lines can mitigate the impact.

3.2. Defining the Pre- and Post-disruption Periods

After finding a suitable disruption instance, a period of at least three weeks to one month before the disruption should be determined as the pre-disruption period. The period after the disruption is divided into intervals of around one month in order to construct the five or six post-disruption periods required for the analysis. The public holidays shall be removed from these periods because the frequency and opening hours of the metro network differs on public holidays compared to normal days (WMATA, 2024d).

3.3. Identification of the Affected Passengers

Once a suitable disruption is identified, the next step is to filter out the passengers who were impacted by it. The approach for identifying affected passengers depends on the type of disruption. If the disruption is a single event lasting several weeks or days, the identification process differs from that used for multiple short disruptions occurring over several days, which last only a few hours or less (e.g. around 30-40 minutes) each. Each disruption type demands a specific method to accurately determine which passengers' travel was disrupted.

3.3.1. First Scenario: One Disruption

Passengers who frequently begin their journeys from the disrupted station are more likely to be affected by the disruption and change their behavior due to their frequent use of the station. Therefore, we identify these frequent travellers as the relevant passengers for our analysis.

To determine frequency, the number of journeys that a passenger has made from the disrupted station within a typical, undisrupted month—potentially during the pre-disruption period is calculated. A threshold is determined to label passengers with journeys greater than that threshold as frequent and thus relevant travellers and others as irrelevant for our study.

Choosing a small threshold would include many passengers who do not frequently use the disrupted stations in the analysis, while a higher threshold would risk excluding some of the relevant passengers. To strike a balance between filtering out irrelevant passengers and retaining information on frequent travellers, we choose four journeys as the suitable threshold.

3.3.2. Second Scenario: Multiple Disruptions

In the second scenario, similar to the first, we begin by filtering out the frequent travellers. These are passengers who, during an undisrupted period, initiated at least four journeys from the affected stations around the time of disruption. For instance, if the disruptions consistently occur around 8 AM in September, we should focus on those who made at least four journeys around 8 AM at that station before September.

To accurately determine whether the start time of a journey aligns with the disruption time, a specific window is defined. This window is the interval (referred to as ”headway”) between the disrupted train's departure time and the departure of the preceding train. An additional thirty-minute margin is added to account for variations, as illustrated in Figure 3.2. We have opted for a 30-minute margin to capture a wider range of tap-in times for passengers. This approach ensures we include those who usually tap in at varying times, not just those with a fixed schedule around the disruption. Using a smaller margin would exclude passengers whose tap-in times differ more broadly, limiting our understanding of the disruption's impact. This wider margin helps us better analyze how different passengers react to the disruption.

Therefore, for each journey i if $tap\ in_i$ falls within the range *[delayed train departure time - headway - 30 min, delayed train departure time + 30 min]* the journey is classified as a frequent journey occurring around the disruption time. This ”headway” parameter is the same as the *headway per line per direction* discussed in Section 3.1.3.

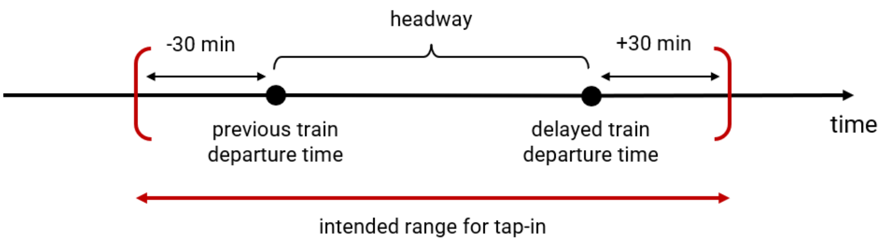

Figure 3.2: The range for $tap\ in_i$ to consider i a frequent journey made around the disruption time.

Having identified the frequent travellers, the next step is to identify the frequent travellers who have experienced the disruptions at least twice. Some passengers might have not experienced any of the disruptions at all. Some might have experienced them only once which might be interpreted as an accidental disruption by the passengers. Therefore, setting the threshold at two disruptions provides evidence that these incidents occurred recurrently, affecting the same passengers multiple times.

The criteria to determine whether journey i of a passenger has been affected by a disruption or not is based on the timing of their tap-in relative to the train's departure. Specifically, a journey is considered affected by a disruption if $tap\ in_i$ falls within the range *[delayed train departure time - headway, delayed train departure time - scheduled headway]*. This range is visualized in Figure 3.3.

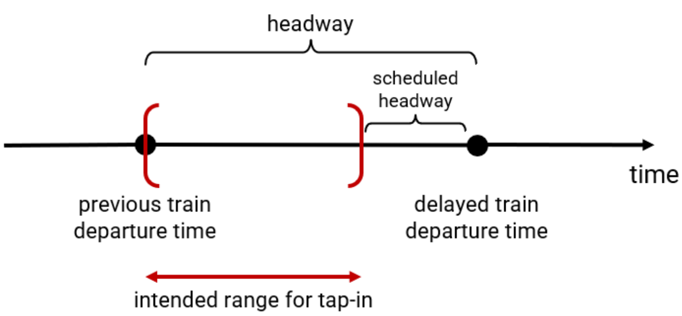

Figure 3.3: The range for $tap\ in_i$ to consider i a journey affected by a disruption (delay).

This criteria ensures that people who have tapped-in just before the departure of the delayed train (and did not fully experience the long delay) are excluded. The scheduled headway is deducted from the departure time for this purpose because it is the usual interval between two trains on that line.

3.4. Identification of the Reference Passengers

To ensure that any observed changes in passenger behavior are directly attributable to disruptions, it's essential to control for external factors such as seasonal trends which can affect travel behavior regardless of disruptions. In order to understand these natural trends we can look at the behavior change of another group of passengers as a reference point.

In the case of a single event disruption leading to a complete station closure, we identify a group of passengers who travel via another station(s) that closely resemble the disrupted station in terms of redundancy, regional characteristics (whether they are in residential or workplace areas), and proximity to the city center. The behavior of these passengers is analyzed as a reference. Otherwise, if we have multiple disruptions leading to delays rather than a complete closure, we can use the users of the disrupted station who have not experienced the disruptions as the reference group.

Importantly, the reference should remain unaffected by any disruptions to accurately reflect the natural trend in passenger behavior. Passengers who traveled via the reference stations for at least four times in the pre-disruption period are selected as the reference group. This criterion ensures that sporadic travelers, including tourists, are excluded from the analysis. Consequently, a temporal travel pattern analysis is conducted for both the affected passengers and those from the reference group. Comparing the differences between the behavior of the affected and reference passengers can provide insights into the impact of disruptions.

3.5. Passenger Segmentation

Using clustering, we can identify unobserved travel patterns that represent passengers' travel behavior across different periods. By tracking how passengers transition between these clusters from one period to the next, we can assess whether changes in behavior align with our expectations, such as whether travel frequency decreases. This approach allows us to draw conclusions about the impact of disruptions. Moreover, investigating each individual's behavior separately is time-consuming and computationally expensive. On the other hand, evaluating passengers at an aggregate level fails to capture their similarities and differences. Clustering provides an effective framework to study the behavioral patterns of different passenger groups. By addressing both their similarities and differences, clustering enables targeted analysis and allows for the tailoring of interventions or policies to meet the needs of different segments.

This section of the study adopts a clustering approach inspired by Briand et al. (2017), which investigated the year-to-year changes in passenger groups, and Eltved et al. (2021), which focused on behavior changes following a disruption using clustering. The current study differentiates itself by evaluating the long-term changes in clusters in several periods, in contrast to the single-event focus of Eltved et al. (2021) and the annual analysis of Briand et al. (2017). Additionally, this research uses the mixture latent Markov model which addresses the limitations of traditional clustering methods such as K-Means which is used by Eltved et al. (2021).

This section begins with an overview of the most widely used clustering methods in travel behavior research in Section 3.5.1. It then introduces the chosen clustering method for this analysis, along with the model assumptions, in Section 3.5.2. Finally, the specific indicators used to capture the travel behavior of passengers are explained in Section 3.5.3.

3.5.1. Clustering Methods

Hierarchical clustering algorithms are designed to create a nested structure of clusters and are categorized into agglomerative and divisive algorithms. Agglomerative algorithms employ a bottom-up approach, and start with considering each data point as a separate cluster. These clusters are then progressively combined based on a measure of dissimilarity until all points are consolidated into one cluster. Conversely, divisive algorithms start with all data points in one cluster and systematically split this cluster until each data point stands alone (Ghaemi et al., 2017). One significant advantage of hierarchical clustering algorithms is their ability to visually represent the clustering results through a dendrogram, which illustrates the sequence of cluster mergers or splits. Additionally, the number of clusters does not need to be predetermined; various numbers of clusters can be achieved by cutting the dendrogram at different levels. However, this method demands substantial computational power and requires a decision on the level at which to cut the dendrogram to obtain the desired clusters (Ran et al., 2023; Xu and Tian, 2015).

K-means is an unsupervised clustering algorithm that groups data points into k different clusters based on their distance to the cluster center. The data points within each cluster are similar to one another but distinct from those in other clusters. The K-means clustering algorithm is valued for its efficiency, flexibility, and low computational complexity, making it particularly suitable for handling large datasets (Ikotun et al., 2023). However, K-means has several drawbacks: it requires pre-specifying the number of clusters, may converge to a local optimum, is sensitive to outliers, and performs poorly with non-convex data (Xu & Tian, 2015).

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is an algorithm that identifies dense clusters in data by separating them from regions of lower density. DBSCAN operates with two parameters, ϵ and $MinPts$. $MinPts$ represents the minimum number of points to form a cluster, with clusters falling below this threshold classified as noise. ϵ indicates the maximum distance between two points to be considered as neighbors. Unlike K-Means, which is influenced by noise and outliers, DBSCAN can automatically handle noise (Ester et al., 1996). Additionally, DBSCAN is capable of detecting clusters in data with arbitrary and non-convex shapes. The downside however, is that the result of the clustering is highly sensitive to the value of the parameters and the outcome of the clustering could be of low quality if the data space's density varies significantly (Xu & Tian, 2015).

Model-based clustering approaches assume that the data is generated by a combination of underlying probability distributions (Magidson & Vermunt, 2002). Each cluster is indicated by a parametric distribution, and the overall data is modeled as a mixture of these distributions (Xu & Tian, 2015). One of the common model-based clustering algorithms is the Gaussian mixture model (GMM). GMM assumes that the data is generated based on a mixture of several Gaussian distributions and the data generated from the same distribution belong to the same cluster (Xu & Tian, 2015). The GMM algorithm however has a relatively high time complexity. The assumption that the data within each cluster is normally distributed may not hold for all datasets and might potentially lead to poor clustering performance. Additionally, GMMs are sensitive to the initial parameter values; poor initialization can lead to convergence at local optima, thereby compromising the quality of the clustering results (Xu & Tian, 2015).

3.5.2. Model conceptualization

Hierarchical clustering methods are computationally expensive for large datasets and require a decision over which level of the dendogram to cut to obtain clusters. K-Means requires the arbitrary specification of the number of clusters and is highly sensitive to outliers. GMM models are time-intensive and their underlying assumption may not always hold.

Therefore, in this study we use one of the variants of a model-based clustering approach called latent class analysis (LCA) which assumes that each data point belongs to one of the K latent or unobserved discrete classes. Data points belong to a similar latent class or cluster based on the similarity of their values regarding a number of observed indicator variables. We use a variant of LCA to cluster passengers' behavior in each period, thus the latent variable represents travel pattern. The main assumption of LCA is that membership in each latent class can explain the response patterns to the indicators. The relationship between the indicators and latent class variables is estimated via multinomial logit models. LCA typically uses maximum likelihood estimation to estimate model parameters. This involves finding the parameter values that maximize the likelihood function, which indicates the probability of the observed data given the model parameters (Magidson & Vermunt, 2002).

One of the main advantages of LCA is that the decision over the optimal number of clusters is guided by statistical tests, making it less arbitrary. Additionally, LCA accommodates variables of mixed scale types without the need for standardization (Magidson & Vermunt, 2002).

Despite the capabilities of LCA, there are a few limitations to this method. LCA assigns individuals to classes based on the probability of class membership given a certain response pattern; therefore, the most suitable class membership is not guaranteed and the share of individuals in each cluster cannot be precisely determined. Moreover, in LCA the latent classes are named by researchers; thus, issues could arise from whether the class labels properly represent the behavior of the class members (Weller et al., 2020).

The Mixture Latent Markov Model

Our objective is to understand how passengers' behavior evolves over time, which requires a longitudinal analysis. That is why we utilize a variant of LCA particularly suited for clustering longitudinal data: the Latent Markov Model (LMM) (Vermunt & Magidson, 2013). This model allows individuals to transition between clusters over time, with these clusters—referred to as latent *states*—representing different travel patterns in our study. The dynamic nature of LMMs offers a significant advantage over static models, as it enables us to capture the evolving nature of travel behavior by estimating the likelihood of passengers transitioning between different travel patterns over time.

A basic LMM assumes that all individuals share the same transition probabilities, meaning each person has an equal likelihood of moving from one travel pattern to another over time. However, in reality, there may be differences in how individuals transition between travel patterns. To account for this variability, we employ a mixture latent Markov model (MLMM), which is a case of LMM that accommodates heterogeneity in transition probabilities by assuming the population is composed of several *classes*, each with its own Markov process and distinct transition probabilities (Kroesen & van Cranenburgh, 2016).

Inspired by the approach of Kroesen and van Cranenburgh (2016), we refer to each class as a "mobility style", reflecting the unique characteristics and underlying attitudes with which passengers change their travel patterns over time. Identifying these varying mobility styles, rather than assuming homogeneity, is important for our analysis as it allows us to observe what mobility styles exist and whether the affected passengers are more or less likely to belong to each of these mobility styles. For instance, if we expect a reduction in travel frequency after a disruption, we can see whether a mobility style reflecting this trend has emerged or not. Additionally, we can assess whether affected passengers are more likely than the reference group to belong to this particular mobility style due to them having experienced the disruption. That is why uncovering the different mobility styles is necessary.

Moreover, MLMMs are well-suited for handling missing data, which is a particular challenge in this study, where some passengers may not have recorded activity during certain periods. These models are also capable of managing a large number of time points and supporting indicators of various scale types, thus offering great flexibility in the types of indicators that can be included (Vermunt & Magidson, 2013).

Another advantage of using MLMMs that motivated our choice is their ability to incorporate additional variables, often sociodemographic factors, known as covariates (Magidson & Vermunt, 2004), such as age, gender, education, etc. Although these variables are not part of the observed indicators, they can still impact transition probabilities, travel patterns, and mobility styles. This capability is crucial for our study because whether a passenger is affected or reference may influence their likelihood of belonging to a specific mobility style and their transition probabilities. Therefore, "type" (i.e., whether a passenger is affected or a reference) is included in the model as a covariate, and its significance is subsequently tested.

One of the strengths of MLMMs is their ability to differentiate between individuals who maintain the same travel pattern over time (referred to as "stayers") and those who change their travel pattern (referred to as "movers"). Separating the stayers allows us to more accurately capture the heterogeneity among the movers (Kroesen & van Cranenburgh, 2016), thereby providing a better understanding of the phenomena of behavioral stability and change within the population. This approach is often referred to as the "mover-stayer" model.

For stayers, the model assigns a probability of 1 for remaining in the same state, indicating no transition in their travel pattern, and a probability of 0 for changing states, meaning their probabilities are fixed at either 0 or 1, with no intermediate values. In contrast, for movers, the model calculates probabilities that describe the likelihood of transitioning from one state to another between consecutive periods, with these probabilities ranging from 0 to 1. This process follows a Markov model, where the current state membership is assumed to depend on the previous state membership (Magidson et al., 2009).

Model Assumptions

First, it is assumed that passengers' travel behavior can be captured by a set of latent variables, i.e. travel patterns, which vary over time and at each period represent travel behavior with the smallest possible number of logit parameters. Six features are used as indicators of these latent variables: share of active days, average number of journeys per weekday, share of weekend journeys, share of peak-hour journeys, share of days with similar first boarding time, and share of days with similar last boarding time (feature selection is discussed in Section 3.5.3).

Secondly, we assume that another latent class variable, i.e., mobility style, explains the initial travel pattern membership and the probabilities of transitioning between different travel patterns over time. Mobility styles capture the heterogeneity in the transition probabilities. Contrary to travel patterns, mobility styles are not dependent on time (Vermunt & Magidson, 2013). For example, one such mobility style mentioned earlier is the stayers who do not change behavior over time. Mobility styles are not known beforehand and are identified after the clustering.

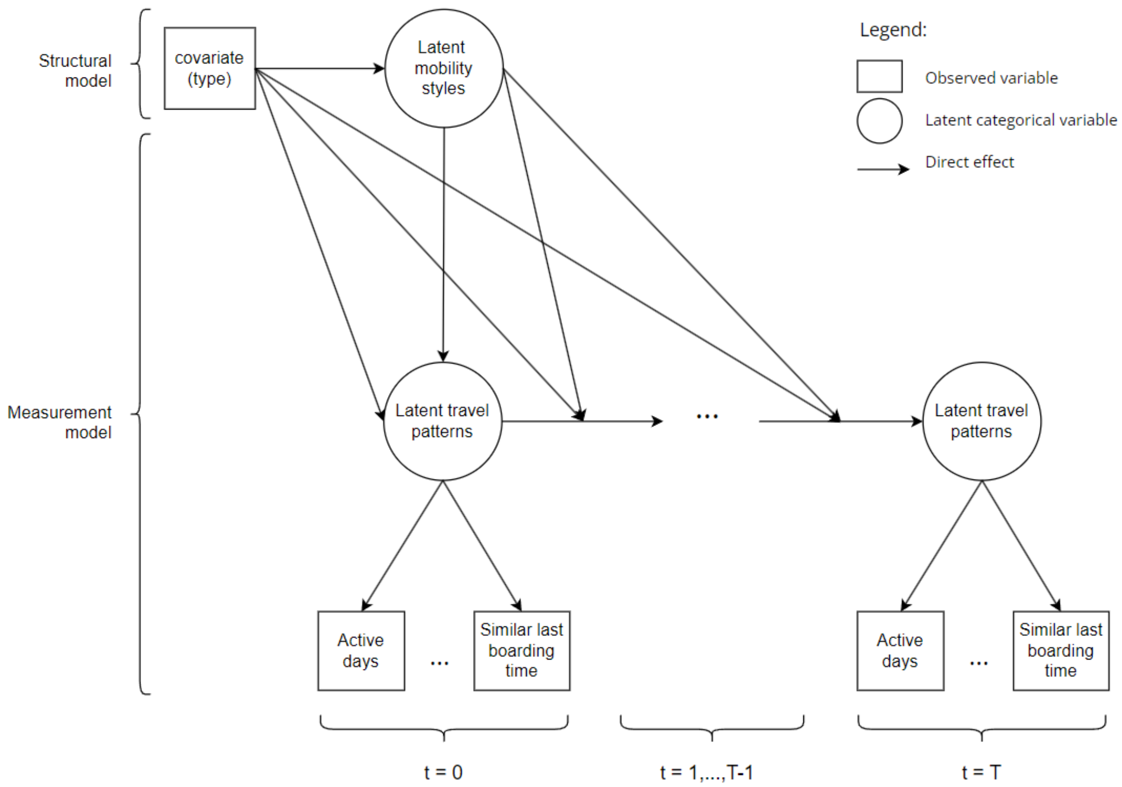

Figure 3.4: Graphical representation of the latent class model.

Third, the data points are clustered based on their similarity regarding the indicators. This part of the model is called the measurement model (Haustein & Kroesen, 2022). In a latent Markov model the indicators are assumed to be independent conditional on a latent process. This assumption is known as the local independence assumption (Bartolucci et al., 2010). This assumption is graphically presented in Figure 3.4, where the six indicators are assumed to be independent conditional on the latent travel patterns.

Fourth, the part of the model which probabilistically assigns data points to the latent classes is called the structural model. This part of the model allows for the inclusion of additional covariates. Figure 3.4 graphically illustrates the structural model where the covariate "type" is included in the model as an influential factor in class membership, initial state membership and the transition probabilities.

Figure 3.4 shows that for every period $t \in \{0, ..., T\}$ the Markov model includes a latent state variable (travel pattern). Given that the latent state variable is nominal, a multinomial logit model is utilized to estimate the relationship between successive latent state variables. The resulting logit parameters can be used to create a matrix of transition probabilities, which indicate the likelihood of latent state membership at the next period based on the current period. Thus the first-order Markov assumption indicates that the probability of transitioning to a future state depends on the current state, and not on any previous states (Vermunt & Magidson, 2013).

Another assumption is that the logit parameters that describe the relationships between the latent travel patterns and the indicators are equal in all periods. This implies that the structure and nature of the travel patterns does not vary over time. This assumption is known as the measurement invariance assumption. This assumption is crucial for interpreting transitions between travel patterns over time because it is difficult to analyze transitions if the patterns in one period are different from the next (Kroesen & van Cranenburgh, 2016).

3.5.3. Selected Indicators

We have specific expectations regarding how passenger behavior may change following a disruption, which inform our selection of indicators for the model. For example, we anticipate a decrease in travel frequency, with frequent travelers making fewer journeys on fewer days. This change could be driven by passengers choosing to work from home or use their own vehicles on certain days after the disruption. Therefore, it is essential to include indicators that effectively capture this expected shift in behavior.

Additionally, the timing of the disruption—whether it occurs during peak or off-peak hours—can affect passenger activity during peak times. This necessitates another indicator to account for this aspect. If the disruption happens on weekdays, passengers might reduce their use of public transport for commuting, opting instead to use it primarily for leisure activities on weekends. Furthermore, to avoid frequently disrupted hours, passengers might alter their usual travel times on certain days, leading to variations in their boarding times. Therefore, indicators are chosen based on these expectations, data availability and limitations, as well as insights from the literature review on passenger segmentation indicators.

Based on these indicators, travelers are clustered, and changes in their behavior due to a disruption are analyzed. The indicators are divided into several categories:

Travel frequency: Travel frequency is a crucial aspect of passenger segmentation, helping to identify passengers who use public transport intensively versus those who use it sporadically. This distinction provides insights into the users with the highest contribution to the total journeys in public transport. The following indicators are used in this study to capture the travel frequency of each passenger during the study period:

- **Share of active days**

  This indicator is defined as the percentage of days during the period on which a passenger made at least one journey. In order to calculate this feature, the number of unique days on which at least one journey has been recorded for a card ID is calculated. This features ranges from zero to one.

- **Average number of journeys per weekday**

  This indicator is defined as the average number of journeys a passenger has made on weekdays (Monday to Friday). In order to calculate this feature, the data is first filtered on the weekdays. Next, the number of journeys per passenger is calculated and divided by the total number of weekdays in the period.

Time-of-day / Day-of-week: Another aspect of the temporal travel pattern is the time-of-day or day-of-week in which passengers start their journeys. This feature provides useful insights into the characteristics of a passenger, e.g., whether the passenger often travels during the peak hours and whether he/she travels mostly during the week rather than the weekend.

- **Share of weekend journeys**

  This indicator is defined as the share of journeys during the whole period which were made on Saturday and Sunday per passenger. In order to calculate this feature the data is filtered on the weekends, the number of weekend journeys per passenger is calculated and divided by the total number of journeys per passenger in the period.

- **Share of peak-hour journeys**

  This indicator is defined as the share of weekday journeys in the whole period which were made during the peak hours per passenger. The definition of peak-hours is derived from WMATA's website where 6AM to 9AM in the morning and 3PM to 7PM in the evening is indicated as the busiest periods of the day (WMATA, 2023). In order to calculate this feature, the data is first filtered on the weekdays. Next, the number of peak-hour journeys per passenger is calculated and divided by the total number of weekday journeys in the period.

Travel regularity: This category of indicators account for the travel pattern regularity and flexibility.

- **Share of similar first boarding time**

  This indicator is defined as the share of days a passenger started his/her first journey of the day around a similar time. This indicator enables the identification of travel time regularity and the extent to which passengers travel at the same time everyday which is a suitable measure of time flexibility of passengers.

  To calculate this indicator, some studies divide each day into bins of one hour or 30 minutes and calculate the number of days a passenger starts his/her first journey in a similar bin. However, one of the limitations of this method is that journeys occurring near the border of the bins might be placed into different bins as is the case for two journeys at 08:02 AM and 07:58 AM and thus be considered as two journeys with different boarding times. Therefore, an alternative approach introduced by Bhaskar, Chung, et al. (2014) is to use DBSCAN to identify the densest areas in the dataset as a group of days with approximately similar boarding times. In order to calculate this indicator, the data is filtered on the first tap-in of each day per passenger. Next, the tap-in time is translated into minutes-from-midnight, e.g., 07:00 AM is transformed into 420 minutes.

  DBSCAN requires two parameters, ϵ and $MinPts$, as input. $MinPts$ represents the minimum number of points to form a cluster, with clusters falling below this threshold classified as noise. Moreover, ϵ indicates the maximum distance between two points to be considered as neighbors. As a rule of thumb, the $MinPts$ is twice the dimension (i.e., two in this case). Since eventually the largest cluster is selected as the value for the number of similar boarding times, the value of $MinPts$ is trivial as smaller clusters are set aside automatically.

  To determine the proper value for ϵ, Sander et al. (1998) argues that the ϵ of the smallest (least dense) clusters are good candidates for this value. Given that $MinPts$ is known the first step is to calculate the distance of each data point to its kth nearest neighbor. Here k is $MinPts$. Then the distances are sorted in descending (or ascending) order and plotted in a graph like Figure 3.5. This graph provides some insights into the density distribution of the data points. The elbow of the graph is chosen as the value for ϵ. All the points with a larger distance than the elbow are labeled as noise, and all the points with a distance smaller than the elbow is assigned to a cluster.

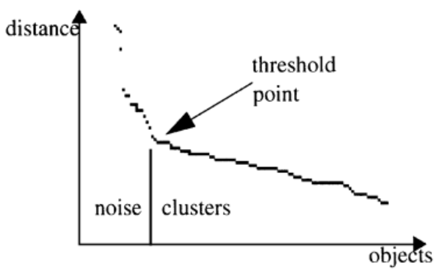

Figure 3.5: Example of a sorted distance plot (Sander et al., 1998).

  The output of this DBSCAN algorithm consists of several clusters, each representing groups of days with similar boarding times per passenger. We use the size of the largest cluster, divided by the total number of active days for a passenger, as the value for this feature. For instance, if the algorithm identifies two clusters for a passenger's boarding times—such as one with 15 days and another with 8 days—the largest cluster size, 15 days, is selected. If the number of active days of the passenger is 30 days, the value for this feature would be 0.5.

- **Share of similar last boarding time**

  This indicator is defined as the share of days a passenger started his/her last journey of the day around a similar time. A similar approach to calculating the first boarding time is used to calculate this indicator.

The indicators are calculated for each affected and reference passenger based on their smart card transactions during each period. This results in a dataset where each row represents the behavior of one passenger, defined by the six specified indicators for a particular period. These rows are then sorted by passenger and ordered chronologically by period, as the Markov algorithm uses the first row for each passenger as the initial period. This dataset is then clustered using the ”Markov” model in the LatentGold 6.0 software package. To avoid local optima, the model is estimated using 200 sets of random starting values for the parameters and 200 iterations for the Expectation-Maximization (EM) algorithm. These settings can be adjusted in the ”Technical” tab of the model.

3.6. Clustering Procedure

In the previous sections, the chosen clustering method, the model assumptions, and the selected indicators were discussed in detail. In this section, the specific steps involved in performing the clustering, and model estimation are outlined.

3.6.1. Model Estimation

The mover-stayer model is composed of several mobility styles (classes) and travel patterns (states). To identify the optimal number of travel patterns and mobility styles, we use the two-step approach introduced by Kroesen and van Cranenburgh (2016). In this method, the number of travel patterns is determined first (without including the latent mobility styles) and is then used to decide the number of mobility styles.

Determining the Number of States

To identify the optimal number of travel patterns (states), we estimate LCA models (without any covariates) starting with one state and incrementally increasing the number of states up to 10. The number of states is input into the model via the ”Variables” tab. The criterion chosen to determine the optimal model is the Bayesian Information Criterion (BIC), which identifies the optimal model based on the model fit and complexity (as measured by the number of parameters) and is typically used for this purpose (Nylund et al., 2007). The model which shows the lowest BIC value is selected as the optimal model, thus determining the optimal number of states.

Determining the Number of Classes

After determining the optimal number of states, a similar approach is used to identify the optimal number of mobility style classes. We begin by estimating mixture latent Markov models with five states and a varying number of mobility styles, ranging from one to four. From the model with two classes onwards, the ”stayer” class is included in the estimation, which can be configured via the ”Advanced” tab. The model which shows the lowest BIC value is then selected as the optimal model.

Determining the Significance of the Indicators and the Covariate

The next step is to test whether the indicators are significant, i.e., to determine whether there is a meaningful relationship between the latent travel patterns and the indicators in the population. Conversely, insignificance indicates that the indicator does not contribute to distinguishing between the travel patterns and can be removed from the model.

One method to test the significance of the indicators and the covariate is the Wald test (Vermunt & Magidson, 2013). This test evaluates whether the parameters associated with each indicator in the multinomial logit models are significantly different from zero. Specifically, the null hypothesis is that all parameters related to an indicator are zero in the population. The alternative hypothesis is that at least one parameter is not zero, indicating a significant effect (Gudicha et al., 2017). The Wald test calculates a statistic for each indicator, and the p-value determines the significance. If the p-value is found to be smaller than 0.05, we need to reject the null hypothesis, and conclude that the indicator is significant. This p-value can be found under the ”parameters” tab in the output of the model.

The final step is to add the covariate to the optimal model and re-estimate it to test the significance of the covariate. If the covariate is found to be insignificant in its relationship with either initial state membership, transition probabilities, or class membership, the insignificant relationship can be removed from the model via the ”Model” tab.

3.7. Interpretation of the Results

Based on the output of the mixture latent Markov model, we can evaluate travel patterns by analyzing the mean values of indicators within each pattern, following the approach commonly used in studies employing latent class analysis. For simplicity, these studies often assign labels to each pattern to facilitate reference.

Furthermore, the model allows us to examine the probabilities of transitioning between travel patterns over time. By comparing these probabilities for affected passengers and a reference group, we can gain insights into their behavioral differences. Additionally, the characteristics of different mobility styles can be derived from their transition probabilities, and labels can be assigned to these styles accordingly. This approach also enables us to determine whether the outcomes observed in our case study align with our initial expectations. A comprehensive discussion of the results and related analyses is provided in Section 5.

4
Case Study Description

This chapter begins with a description of the metro network used in the case study and the different datasets. Section 4.1 provides information on the disruption identified as the case study. Section 4.2 and Section 4.3 detail the identification of the affected passengers and the reference passengers respectively. Finally, Section 4.4 discusses the preparation of the data for the clustering.

The data which is used in this research is provided by the Washington Metropolitan Area Transit Authority (WMATA) through the Smart Public Transport Lab at TU Delft. The metro network of Washington D.C., also called the Metrorail, consists of 128 miles of track and serves 98 stations across Virginia, Maryland, and the District of Columbia through six color-coded rail lines as seen in Figure 4.1. Metrorail offers service to over 600,000 passengers daily across the Washington, DC area and is the second busiest transit system in the United States. The design of the system ensures that passengers tap-in their smart cards upon entrance to the network as well as the exit (WMATA, 2024c)(WMATA, 2024b).

The data provided consists of comprehensive information on Metrorail's operations and service disruptions from August 2019 to December 2022 which includes Automated Fare Collection (AFC), Automatic Vehicle Location (AVL), planned and unplanned disruptions log.

The AFC dataset provides crucial information on the journeys of individual smart card users, including date, card ID, tap-in time, tap-out time, tap-in station, tap-out station, and train line. For the purposes of this study, we assume each smart card represents a single passenger.

The AVL dataset offers comprehensive data on train movements, with each row representing a train's service at a station, also known as a stop visit. Key fields in this dataset include date, train ID, station name, next station name, arrival time, departure time, train line, direction, headway, and scheduled headway.

The planned disruptions log details each recorded disruption, specifying the date, start time, train line, and description. Similarly, the unplanned disruptions log records each incident's start time, train line, station name, direction, description, passenger delay, line delay, and train delay.

Given that the COVID-19 pandemic began in March 2020 in Washington D.C., leading to a significant decline in public transport usage (WMATA, 2024a), our study focuses exclusively on the pre-pandemic period from August 2019 to March 2020.

4.1. Disruption Identification

In this section the analysis steps and the subsequent results of exploring the planned and the unplanned disruptions log are explained. After mentioning the findings and limitations of these two datasets, the analysis of the AVL data is discussed in detail. At the end, the disruption instance which was chosen to focus on is introduced.

Figure 4.1: Washington DC metro map in 2019

4.1.1. Finding a Disruption from the Planned/Unplanned Disruptions Log

The investigation of the planned disruptions log, as the first source of data to be analyzed, indicated that all of the recorded planned disruptions take place either in the weekend or after 9PM on weekdays. This does not satisfy the first criteria of choosing the suitable disruption. This means that public transport users who travel during the week and in working hours are barely affected by these disruptions.

Additionally, the official website of WMATA along with news about the metro system were explored in order to confirm if there were no major planned maintenance works in the system in the period of interest. The investigation indicated that several stations of the yellow and the blue line were closed for maintenance work from May 2019 until September 2019; however, due to the unavailability of the data, this disruption instance could not be further investigated.

Alternatively, we explored the unplanned disruption log for the period of 21 August to 1 October 2019. Three values are recorded for each logged disruption in this dataset which indicate the amount of delay (in minutes) caused for the passengers, line, and the train. These three values could serve as a proxy of the immediate impact of the disruption.

Figure 4.2 shows the cumulative proportion of delays caused by the disruptions. The figure indicates that over 90% of the disruptions cause delays of less than 10 minutes for either of the passengers, the line, or the train. Moreover, the largest delays are below 50 minutes. Therefore, there are no records of disruptions that last for hours or days. An alternative to large disruptions could be a number of considerable delays (for example above 20 minutes) happening in a short period in the same place. This instance could also have a long term impact on passenger travel behavior. Thus, the next step is to look for such instance.

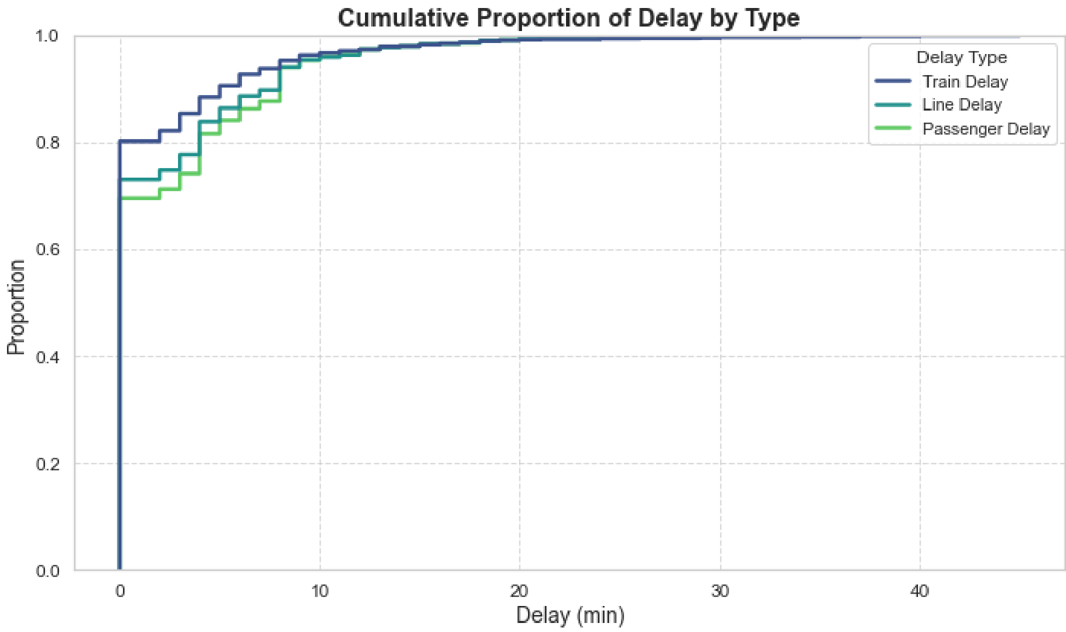

Figure 4.2: Empirical cumulative distribution function plot of the delays.

Although the three measures of delay are available per disruption in the unplanned disruptions log, there are two limitations to this dataset:

(1) First, it does not contain information on the scope of each disruption. For example, it is not possible to infer from this dataset whether a delay in a station propagated to adjacent stations or not.
(2) Second, the exact time when each disruption has ended is not recorded in the dataset.

Thus, we continue the investigation with the AVL dataset as a source of information about the train movements enabling us to identify the scope and duration of each disruption.

4.1.2. Finding a Disruption from the AVL Data

Prior to this analysis, the AVL data was processed and several errors were removed from this dataset:

(1) In this dataset the line name is inferred based on the observed stop sequence. The rows where the line name has not been inferred correctly were removed.
(2) Every station is served by a specific line or a number of lines. For example station 'Glenmont' is served solely by the red line. However, there are instances in the dataset where a station which belongs to a specific line, is served by a train which belongs to another line. Such instances could occur due to vehicle repositioning among other things. Therefore, they were removed from the dataset.

Next, the two measures of *headway per line per direction* and the *headway per direction* were calculated and added to the dataset as indicators of disruptions. Whenever large values are observed for both of these measures, a large delay has happened. Figure 4.3 shows the cumulative proportion of the headways observed for the trains in the period of 21 August to 1 October 2019. It is evident from the graph that the majority of the headways are below 20 minutes which is in line with the findings from the unplanned disruptions dataset.

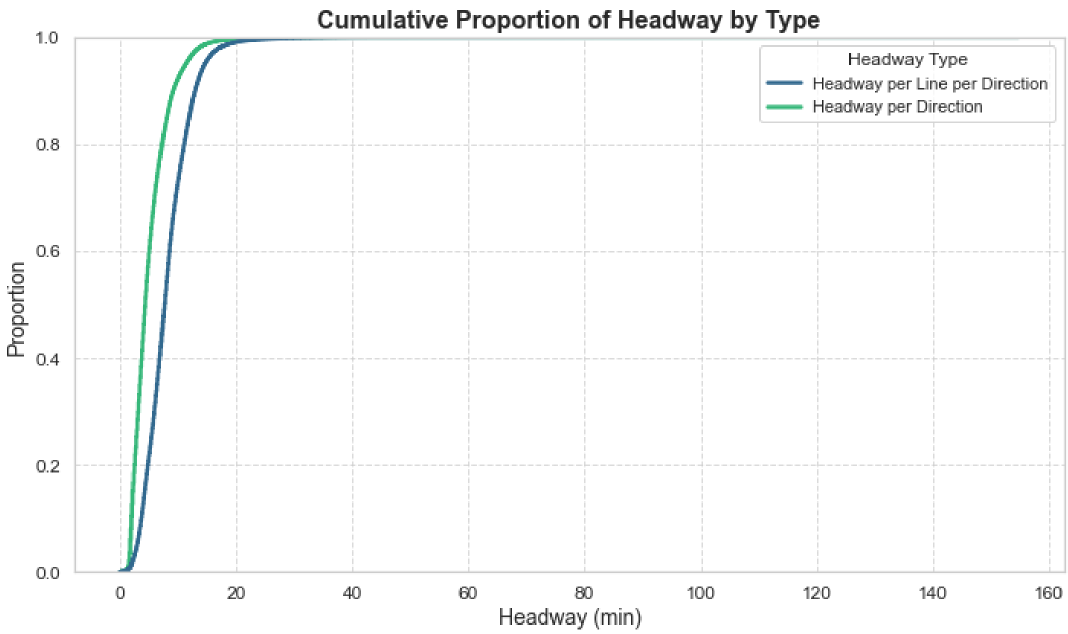

Figure 4.3: Empirical cumulative distribution function plot of the headway.

For every station and every direction, the instances where the values of *headway per line per direction* and *headway per direction* were higher than 20 minutes were filtered to be analyzed. We looked for a short period of time in which multiple large headways happened repeatedly, as this could serve as a suitable instance for the disruption.

4.1.3. The Chosen Disruption Instance

This investigation identified a period in September during which several delays ranging from 20 to 37 minutes affected five stations on the orange line heading towards Vienna. Figure 4.4 shows the locations of the affected stations on the metro network. Table 4.1 provides the specific dates of the disruptions at each station, as well as the total number of delay occurrences per station. Notably, there were days when more than one significant delay occurred. This particular instance meets all the criteria established for selecting an appropriate disruption for our study.

Table 4.1: Disrupted stations and disruption dates.

Disrupted Station	Dates (in September 2019)	Number of Occurrences	Range
Minnesota Ave	9 - 11 - 19 - 20 - 23 - 24	11 times	20 to 37 min
Deanwood	11 - 19 - 20 - 23 - 24	8 times	20 to 33 min
Cheverly	11 - 19 - 20 - 23 - 24	9 times	20 to 32 min
Landover	19 - 20 - 23 - 24	7 times	20 to 32 min
New Carrollton	11 - 16 - 17 - 19 - 20 - 23 - 24	12 times	20 to 31 min

Next, the pre- and the post-disruption periods were determined as can be seen in Table 4.2. These periods were chosen in such a way to keep the ratio of the number of weekend days to the total number of days similar across all of the periods. Moreover, the public holidays were removed from the analysis for all of the periods.

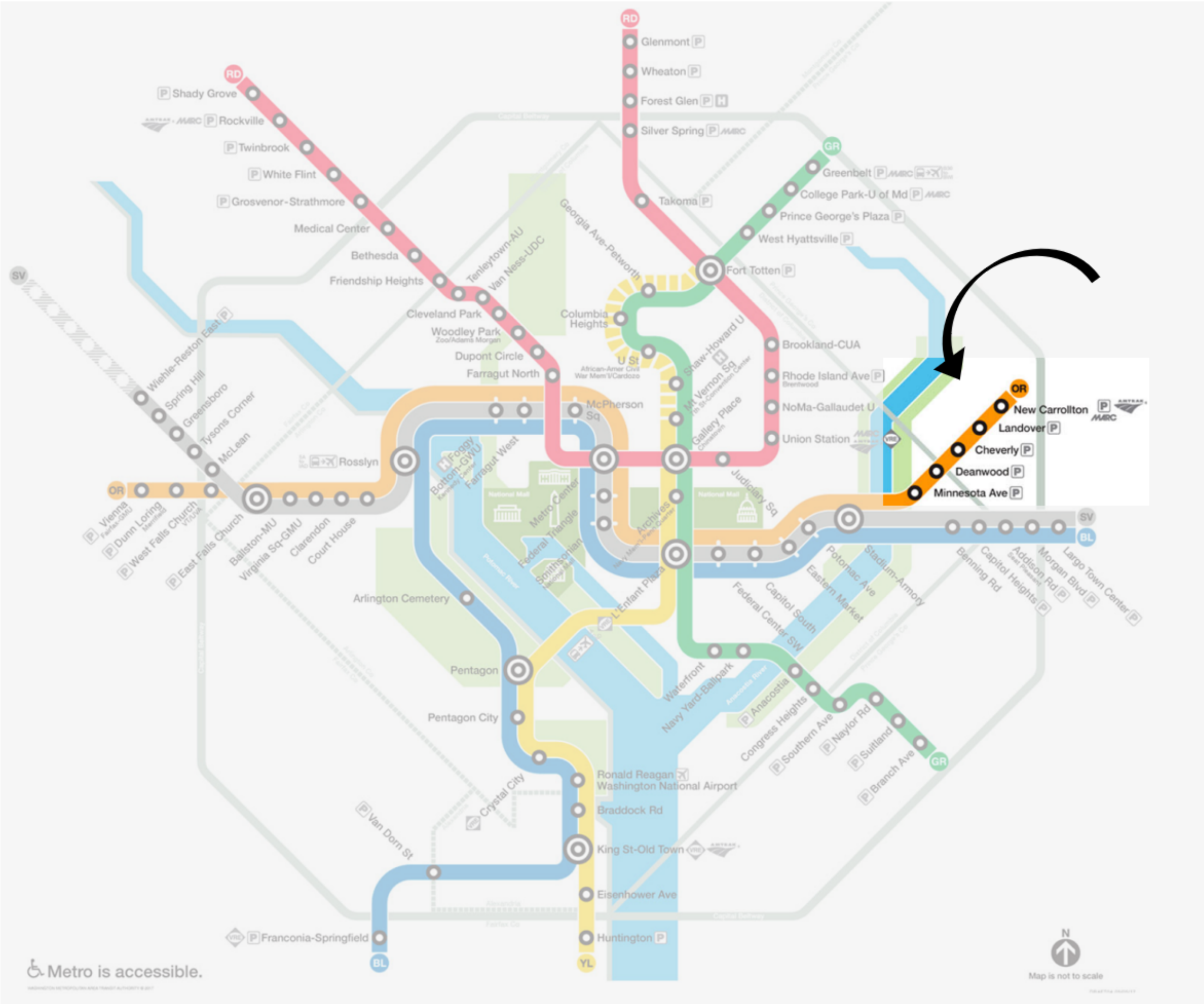

Figure 4.4: Disrupted area identified for the analysis.

Table 4.2: Pre- and post-disruption periods.

Period	Start	End	Duration
Pre-Disruption	2019-08-11	2019-09-07	27 days
Post-Disruption 1	2019-09-25	2019-10-26	31 days
Post-Disruption 2	2019-10-27	2019-11-26	30 days
Post-Disruption 3	2019-11-27	2019-12-21	24 days
Post-Disruption 4	2020-01-01	2020-02-01	30 days
Post-Disruption 5	2020-02-02	2020-02-29	29 days

4.2. Identification of the Affected Passengers

Having identified a suitable disruption to focus on, the passengers who were affected by the disruption were identified and filtered. Since our disruption instance consists of multiple disruptions, we followed the procedure laid out for this specific kind of disruption.

First the people who frequently traveled from the disrupted stations in the pre-disruption period around the disruption time were identified and labeled as frequent travellers. For each journey i if $tap\ in_i$ falls within *[train departure time - headway - 30 min, train departure time + 30 min]* the journey is counted as a journey that has happened around the disruption time. A passenger with at least four of such journeys is a frequent traveller. The total number of frequent travellers for the five disrupted stations with the mentioned criteria amounted to 5976 passengers.

Next, the frequent travellers who had experienced the disruptions at least twice were filtered. If $tap\ in_i$ of journey i of a passenger is in the range *[train departure time - headway, train departure time - scheduled headway]*, that journey is considered to have been affected by the disruption. The total number of affected passengers among the frequent travellers who have experienced the disruptions at least twice amounted to 138 passengers.

4.3. Identification of the Reference Passengers

One of the most suitable candidates for the reference passengers are the frequent travellers who travel via the disrupted station but have not experienced any of the disruptions. Therefore, the identification of the reference can be done in parallel to identification of the affected passengers. After identifying the frequent passengers, i.e. passengers who have started their journeys at least four times from the disrupted stations in the pre-disruption period, we can calculate the number of times each of these frequent passengers have actually experienced any of the disruptions. If a passengers has not been affected by any of the disruptions, he/she can be categorized as a reference passenger. Otherwise, if a passengers has experienced at least two of the disruptions, he/she is labeled as an affected passenger.

The total number of candidates for the reference passengers amounted to 4969 passengers. A last filter is applied here on these candidates to further refine the selection of the reference: Each of the affected passengers is a frequent user of one of the five disrupted stations, e.g., because they live in the vicinity of that stations, and they often use that station. Therefore, we can identify the ”most frequently used station” for each of the affected passengers as well as the reference passengers. Then use the distribution of the stations among the affected passengers, to select the reference passenger.

the distribution of the five stations among the affected passengers in this case study is New Carrollton 60%, Minnesota Ave 14%, Landover 14%, Deanwood 7%, and Cheverly 5%. Thus, we select the reference passengers in such a way to keep the same distribution of stations among them. Eventually, the number of reference passengers amounts to 3965 people.

4.4. Data Preparation for the Clustering

Next, the travel behavior of the affected and reference passengers in each of the six periods (one pre- and five post-disruption) are formed based on calculating the six behavioral indicators using smart card transactions. However first, the AFC data is processed and cleaned using the following filters:

(1) The transactions where the tap-in and tap-out stations are the same are removed.
(2) The transactions where the tap-in and tap-out stations are the same but are presented with different names are removed.
(3) The transactions with missing tap-in and/or tap-out information are removed.
(4) The transactions where the tap-out time is earlier than the tap-in time are removed.
(5) The transactions with journey times longer than three hours are removed. Because the longest paths in the network are around two hours and several minutes.
(6) The transactions with journey times shorter than two minutes are removed. Because the shortest paths in the network are around two minutes.
(7) The duplicated transactions are removed.

This filtering removes 1% of the transactions of the affected passengers (aggregated over all periods) and results in 21090 rows of transaction for the 138 affected passengers. Similarly, the filtering results in a 2% reduction of the transactions of the 3965 reference passengers and consequently 423088 rows of data are left.

After the cleaning process, we calculate the indicators for each passenger per period. This produces a dataset where each row represents the behavior of an individual, characterized by the six indicators for a given period. This dataset is then clustered using the ”Markov” model in the LatentGold 6.0 software package.

5 Results

This chapter presents the results of the analysis and is divided into two main sections. Section 5.1 explains the initial clustering results and how we select the optimal model, based on the ideal number of states and classes. Section 5.2 then examines this optimal model by explaining the resulting travel patterns (states), the transition probabilities between these patterns over time, and the nature of the behavior represented by each identified mobility style. Finally, we compare the observed outcomes with our expected outcomes to assess whether the disruption had the anticipated impact.

5.1. Model Definition

This section begins by outlining the specifications used in calculating the indicators in Section 5.1.1. Following that, Section 5.1.2 discusses the results of the steps taken to determine the optimal model, including the identification of the optimal number of states and classes, along with the reasoning behind these decisions. Finally, Section 5.1.3 examines the significance of the indicators and the covariate, as well as the resulting implications. The purpose of this section is to provide a comprehensive overview of the process involved in obtaining the optimal model.

5.1.1. Calculation of the Indicators

The value of the six indicators are calculated for each of the affected passengers and the reference passengers based on their smart card transactions in the pre- and post disruption periods. Since there are five post-disruption periods and one pre-disruption period, there are six rows calculated for each passenger and stored in one dataset. An illustrative example of the dataset ready to be clustered can be seen in Appendix B.1.

The calculation of the last two indicators, the similar first boarding time and the similar last boarding time using DBSCAN required the specification of two parameters namely ϵ and $MinPts$. Since the boarding time is one dimensional, the value of $MinPts$ is equal to two ($2 \times Dimension = 2$).

As explained in Section 3.5.3, the ϵ of the smallest (least dense) clusters are good candidates for this value. The histogram of the distance of each data point to its kth nearest neighbor (k = $MinPts$) is illustrated in Figure 5.1 for the first boarding times of the affected passenger in the pre-disruption period. The distance of 5 minutes is chosen as the proper value for ϵ. All the points with a larger distance to their kth nearest neighbor than 5 min are labeled as noise, and all the points with a distance smaller than 5 minutes are clustered.

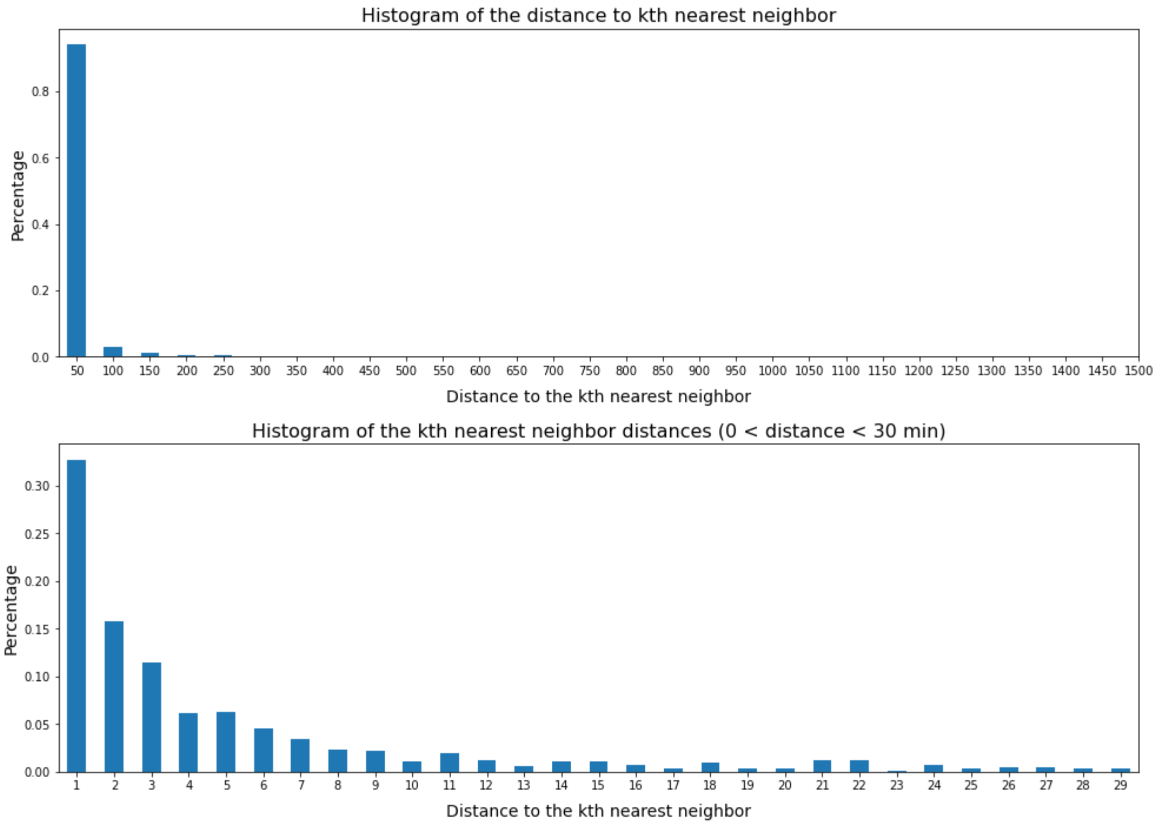

Figure 5.1: Histogram of the kth nearest neighbor distances for the first boarding time

The same approach regarding the last boarding times results in Figure 5.2. The distance of 7 minutes is chosen as the proper value for ϵ for the last boarding times of the affected passenger in the pre-disruption period. All the points with a larger distance to their kth nearest neighbor than 7 minutes are labeled as noise, and all the points with a distance smaller than 7 min are clustered.

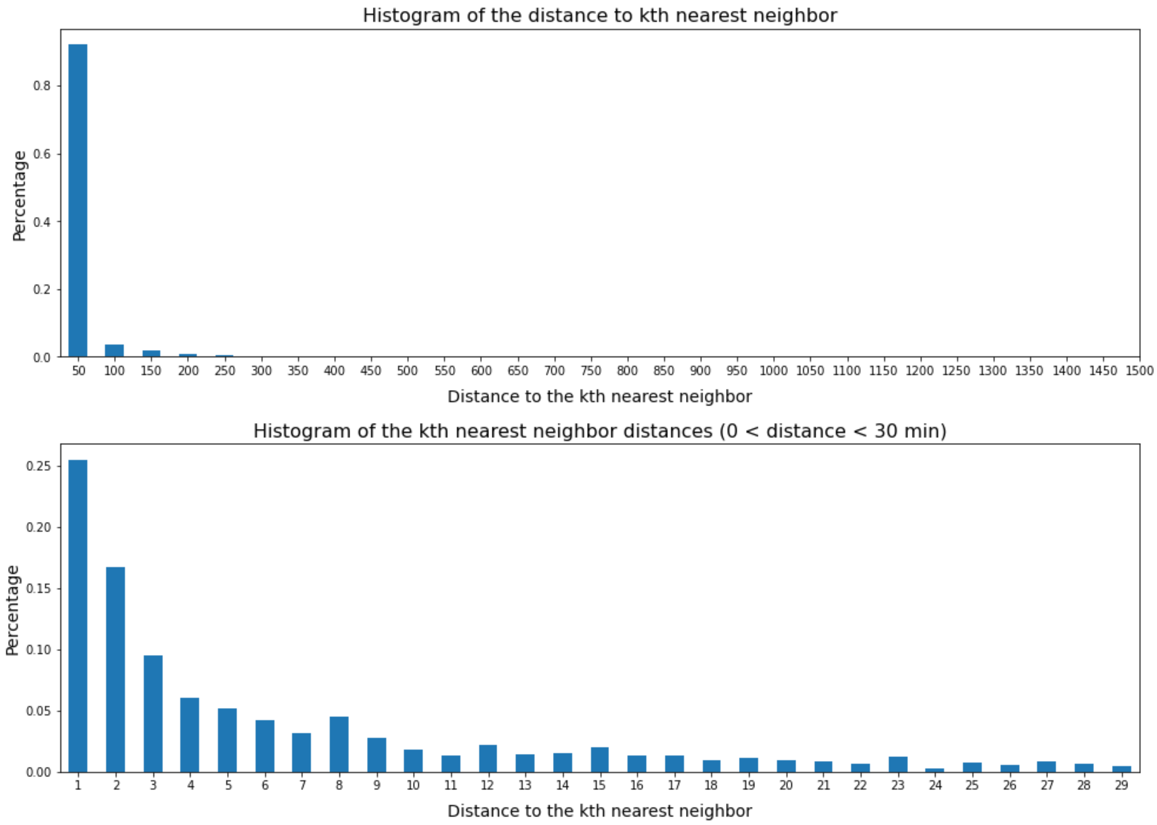

Figure 5.2: Histogram of the kth nearest neighbor distances for the last boarding time

5.1.2. Model Selection and Clustering

To determine the optimal number of travel patterns, we start from estimating the LCA model with one travel pattern and increase the number of travel patterns incrementally up to 10. Typically, the model with the lowest BIC value is selected as the optimal model. However, as shown in Table 5.1, the BIC value tends to decrease consistently as the number of travel patterns increases favoring a model with at least ten states. Since ten states is too large for a mixture latent Markov model to handle (Kroesen & van Cranenburgh, 2016), an alternative approach is used to decide on the best model fit.

To address this, we consider the interpretability and relevance of the clusters as additional criteria for selecting the optimal model. Using these criteria, we chose the model with five clusters because from the six-cluster model onwards the distinction and relevance of clusters decreases and there is not much added value in increasing the clusters.

Table 5.1: Model fit results.

No. of states (travel patterns)	LL	BIC(LL)	Param
1	-12973	26068	12
2	45460	-90667	25
3	57297	-114209	38
4	62975	-125434	51
5	67774	-134901	64
6	72203	-143628	77
7	75294	-149679	90
8	77585	-154129	103
9	79615	-158058	116
10	81721	-162139	129

LL = log-likelihood
BIC(LL) = Bayesian information criterion (based on log-likelihood)
Param = number of parameters
$BIC_{LL} = -2LL + \ln(samplesize)Param$

After the optimal number of states (travel patterns) are identified, a similar approach is taken to determine the optimal number of classes (mobility styles). We start by estimating a mixture latent Markov model with five states and one class and then incrementally increase the number of classes while keeping the number of states constant at five. The models with a ”stayer” class are estimated from two classes onwards. Based on the results illustrated in Table 5.2, the model with the lowest BIC value is the ”1 stayer class 3 mover classes”. However, the decrease in BIC for this model compared to the ”1 stayer class 2 mover classes” is not very substantial while the increase in the number of its parameters (21%) is quite significant. Therefore, we decide to interpret the results using the ”1 stayer class 2 mover classes” model. Next, we add the ”type” as a covariate to the model and re-estimate it.

Table 5.2: Model fit results.

No. of classes (mobility styles)	LL	BIC(LL)	Param
1 mover class	73801	-146904	84
2 mover classes	74488	-148069	109
1 stayer class 1 mover class	74200	-147660	89
3 mover classes	74655	-148195	134
1 stayer class 2 mover classes	74626	-148303	115
4 mover classes	74752	-148182	159
1 stayer class 3 mover classes	74760	-148363	139

LL = log-likelihood
BIC(LL) = Bayesian information criterion (based on log-likelihood)
Param = number of parameters
$BIC_{LL} = -2LL + \ln(samplesize)Param$

5.1.3. Significance of the Indicators and the Covariate

The significance of the indicators are evaluated to determine whether there is a meaningful relationship between the latent travel patterns and the indicators in the population. Conversely, insignificance indicates that the indicator does not contribute to distinguishing between the travel patterns and can be removed from the model.

As mentioned in the methodology section, one of the methods to test the significance of indicators is the Wald test. The Wald test calculates a statistic for each indicator, and the p-value determines the significance. All six indicators in our model have p-values below 0.05 and are thus significant. This finding supports the inclusion of these indicators in the model, as they play a crucial role in differentiating the latent travel patterns.

Next, we add the covariate ”type” to the optimal model and estimate it to evaluate the significance of the covariate. While the results show that the relationship between the covariate and both the initial state membership and transition probabilities is significant (p-value $<$ 0.05), this relationship was not significant for class membership (p-value (0.27) $>$ 0.05). This means that whether a passenger is affected or reference would not change their likelihood of belonging to a certain mobility style. For example, the reference passengers would not be more probable to belong to the stayer class than the affected. Therefore, the relationship between the covariate and class membership is removed from the model. The estimated model parameters are displayed in Appendix B.2, and are used for the calculation of the results that are provided in the following section.

5.2. Analysis of the Results

In this section we discuss the optimal model in detail. First, Section 5.2.1 explains what behavior each of the resulting travel patterns exhibit and provides a description and label for each. Next, Section 5.2.2 shows the overall probabilities of transitioning from each pattern to the other over time for the affected and reference passengers. Here we can observe how probable passengers are to not change their behavior (inertia) or if they do change behavior to what direction and to what extent. Therefore, we are able to compare the observed outcome with what we expected of behavior change directions. Next in Section 5.2.3, we analyze the transitions in more detail by looking into the mobility styles which differentiate between the individuals based on their transitions. The characteristics of each mobility style is explained followed by a comparison between the affected and reference passengers regarding these styles in Section 5.2.3.

5.2.1. Travel Pattern Profiles

The travel pattern sizes (averaged over periods) and the mean value of the indicators for each travel pattern are displayed in Table 5.3. In addition, Figure 5.3 displays the range and different quartiles for each indicator. The information is displayed for the whole data, thus both the affected and reference. Using this information we describe the characteristics of each travel pattern and provide a representative label for them. However, first, we mention the criteria that we used for naming the patterns:

The patterns with a share of active days above 50% and journeys per weekday above one are called *frequent* travellers because the large values of these two indicators shows a high travel frequency. The patterns with a share of active days below 50% and journeys per weekday below one are called *occasional* travellers because the small values of these two indicators show a low travel frequency. Lastly, those with a value below 10% for the share of active days and journeys per weekday are considered as *sporadic* traveler because these passengers rarely travel with metro. The patterns who performed more than 50% of their journeys during peak hours are called *peak* travellers. Finally, patterns with values below 50% for the share of days with similar first boarding time and values below 40% for the share of days with similar last boarding time are called *flexible* travelers.

Table 5.3: Profiles of the latent travel patterns (for the affected and reference passengers, averaged over all periods).

	1. Frequent less flexible peak travellers	2. Occasional less flexible peak travellers	3. Frequent flexible peak travellers (with weekend)	4. Occasional very flexible travellers (with weekend)	5. Sporadic travellers
Cluster size	33%	27%	19%	15%	6%
Indicators (mean)					
share of active days	0.57	0.25	0.70	0.33	0.04
journeys per weekday	1.53	0.62	1.55	0.55	0.07
share of weekend journeys	0.00	0.00	0.15	0.28	0.16
share of peak journeys	0.76	0.63	0.59	0.46	0.40
share of similar first boardings	0.56	0.47	0.44	0.29	1.00
share of similar last boardings	0.49	0.44	0.38	0.29	1.00

- pattern 1 (size: 33%): **Frequent less flexible peak travellers**

  These passengers are classified as frequent travelers due to their high travel frequency, being active on 57% of days and averaging 1.53 journeys per weekday. They do not travel on weekends and have the highest proportion of journeys during peak hours (76%). Because their peak-hour activity exceeds 50%, we refer to them as peak travelers. Additionally, their boarding times show moderate flexibility compared to other groups, with 56% of days having a similar first boarding time and 49% having a similar last boarding time.
- pattern 2 (size: 27%): **Occasional less flexible peak travellers**

  These passengers are the second least active group, traveling on only 25% of days with an average of 0.62 journeys per weekday. Hence, they are referred to as occasional travelers. They show no interest in traveling on weekends and make 63% of their journeys during peak hours. Their boarding times indicate moderate flexibility, with 47% of days having a similar first boarding time and 44% having a similar last boarding time.
- pattern 3 (size: 19%): **Frequent flexible peak travellers (with weekend)**

  These passengers are the most active group, traveling on 70% of days with an average of 1.55 journeys per weekday. They show an interest in weekend travel, accounting for 15% of their journeys, and 59% of their trips occur during peak hours. This group is the second most flexible in terms of boarding times, with 44% of days having a similar first boarding time and 38% of days having a similar last boarding time.
- pattern 4 (size: 15%): **Occasional very flexible travellers (with weekend)**

  These passengers are active on only 33% of the days with 0.55 journeys per weekday and are thus called occasional travelers. They perform 28% of their journeys in the weekend and are the most active pattern in the weekend. They perform 46% of their journeys during peak hours and are the second least active pattern during peak hours. They are the most flexible group with similar first and last boarding times of 29%.

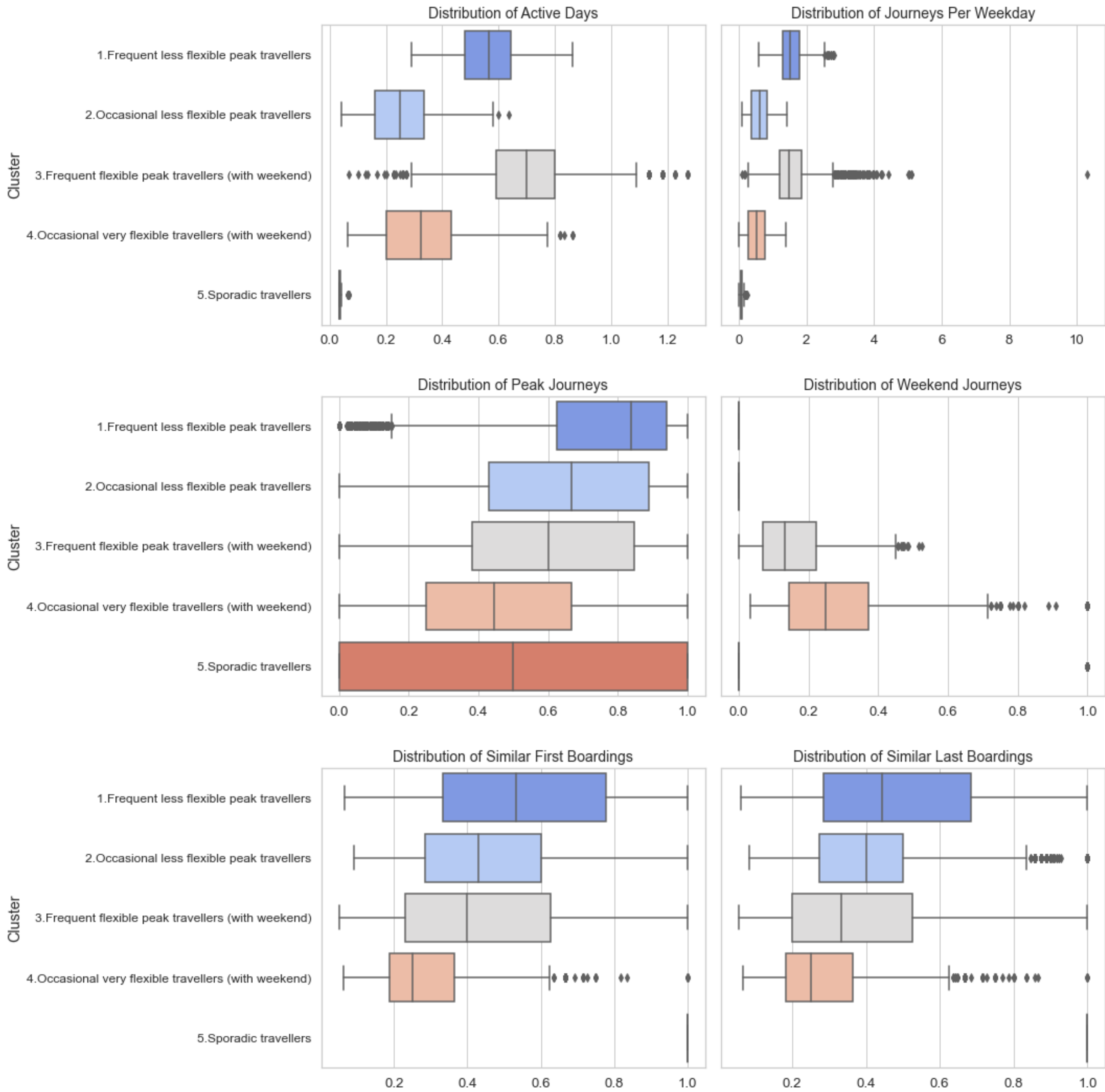

Figure 5.3: Box plots of each indicator and each travel pattern.

- pattern 5 (size: 6%): **Sporadic travellers**

  These passengers are active on just 4% of days, with an average of 0.07 journeys per weekday. Due to their infrequent travel, they are referred to as sporadic travelers. They make 16% of their journeys on weekends and 40% during peak hours. Additionally, they tend to start their first and last journeys at similar times each day.

5.2.2. Travel Pattern Transition Probabilities

Transition probabilities represent the likelihood of individuals shifting from one travel pattern to another between consecutive periods. By analyzing these probabilities, we can determine the extent and direction of passengers' behavior changes, allowing us to assess whether these shifts align with our initial expectations and whether the disruption had a significant impact. Additionally, these probabilities reveal the degree to which passengers maintain consistent travel behavior, known as travel behavior inertia. Overall, this analysis helps us evaluate the effects of the disruption on passenger behavior and understand the dynamics of these changes.

We can examine whether these transition probabilities differ across the five pairs of consecutive periods. However, interpreting five separate transition matrices is complex, and the number of model parameters increases significantly. To simplify the model, we assume that transitions among the post-disruption periods are consistent, reducing the analysis to two matrices: one for the transition from the pre-disruption period to the first post-disruption period, and another for transitions between consecutive post-disruption periods.

To test this, we introduce a dummy variable into the data, assigning a value of 0 for pre-disruption period rows and 1 for all post-disruption period rows. This dummy variable is then added as a time-varying covariate and the model is re-estimated. However, this model fails to converge to an optimal solution even after 200 iterations and 200 parameter starting sets. This suggests that the model was unable to identify an appropriate set of parameters to accurately capture the model relationships, indicating potential issues with model specification. To address this, we may need to simplify the model by reducing the number of parameters and relaxing certain constraints. Another approach could be to increase the number of iterations and starting sets, although we have already set this number quite high (200) compared to the default. As a result, we opt to relax the constraint of varying transition probabilities and instead assume that these probabilities are consistent across all pairs of consecutive periods which might be because the disruption did not lead to a substantial difference in the transitions between periods in the first place.

Table 5.4 shows the overall transition probabilities of the affected and reference passengers. The values which are substantially large (greater than 0.15) have been highlighted in bold. Below we provide a discussion about our observations based on this table, the expected outcomes and the actual observed outcomes. We interpret these observations to derive conclusions about the behavior changes of the affected passengers and their difference with the reference.

Table 5.4: Transition matrices of the affected and reference passengers.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.849**	0.064	0.083	0.003	0.000
	2. Occasional less flexible peak travellers	**0.179**	**0.675**	0.018	0.095	0.035
	3. Frequent flexible peak travellers (weekend)	**0.235**	0.011	**0.656**	0.080	0.019
	4. Occasional very flexible travellers (weekend)	0.113	**0.367**	0.138	**0.356**	0.026
	5. Sporadic travellers	0.012	**0.609**	0.012	0.013	**0.355**
Reference	1. Frequent less flexible peak travellers	**0.772**	0.121	0.089	0.009	0.009
	2. Occasional less flexible peak travellers	0.091	**0.699**	0.026	0.094	0.091
	3. Frequent flexible peak travellers (weekend)	**0.174**	0.081	**0.618**	0.111	0.016
	4. Occasional very flexible travellers (weekend)	0.006	**0.234**	0.048	**0.612**	0.100
	5. Sporadic travellers	0.021	**0.292**	0.018	**0.165**	**0.504**

Travel Behavior Inertia

This table shows that passengers in the three largest travel patterns (1.frequent less flexible peak travellers, 2.occasional less flexible peak travellers, and 3.frequent flexible peak travellers (weekend)) tend to remain in the same pattern, with a likelihood of at least 66% among the affected group and 62% among the reference group. These three patterns contribute to almost 80% of the data (based on Table 5.3). This large likelihood of remaining in the same travel pattern from each period to the next is an indication of travel behavior inertia among the majority of the passengers.

The highest probability of staying in the same travel pattern is observed among passengers in the largest pattern, pattern 1 (frequent less flexible peak travelers). These passengers have an 85% probability of maintaining their behavior from one period to the next in the affected group and a 77% probability in the reference group. This is significant given that pattern 1 is the largest in both groups.

Another insight is that among the affected group, the frequent travelers (patterns 1 and 3) are more likely to stay in the same behavioral pattern (85% and 66%, respectively) compared to occasional travelers, i.e., patterns 2 and 4 (67% and 36%, respectively). This indicates a strong travel behavior inertia among passengers with high travel frequency. This observation aligns with the findings of Adelé et al. (2019) and Li, Yao, Yamamoto, Tang, and Liu (2020), which noted that frequent travelers are less likely to change their behavior and use other modes when facing disruptions. While these studies analyzed the immediate response of passengers to disruptions, our findings suggest that travel behavior inertia is also observed over a longer period following the disruption.

Travel Behavior Inertia: Affected vs. Reference

We expected to observe lower travel behavior inertia (likelihood of staying in the same travel pattern) among the affected passengers compared to the reference group. This expectation was based on the idea that affected passengers, having experienced disruptions, would be more likely to modify their travel behavior. However, this expectation is not fully met. While the diagonal values—indicating the likelihood of staying in the same pattern—are indeed larger among the reference group compared to the affected passengers for patterns 2, 4, and 5, this trend does not hold for patterns 1 and 3. In these patterns, the affected passengers exhibit higher diagonal values, indicating greater inertia than the reference.

Travel Frequency

We initially expected that disruptions might lead passengers to travel less frequently with public transport, possibly opting to work from home or use other modes of transportation like driving a car. Therefore, we anticipated high transition probabilities from frequent to occasional travel patterns among the affected group. However, based on Table 5.4, the data shows a different trend. Among the affected passengers, the majority of frequent travelers (patterns 1 and 3) remain in the same pattern, with probabilities of 85% and 66%, respectively. Additionally, 23% of passengers in pattern 3 shift to pattern 1, which is also a frequent travel pattern.

The behavior of the affected occasional travelers (patterns 2 and 4) also contradicts our expectations. Among passengers in pattern 2, 20% transition to patterns 1 and 3, indicating a shift towards more frequent travel behavior. Similarly, for passengers in pattern 4, 25% transition to patterns 1 and 3 (frequent travelers). Lastly, among sporadic travelers who rarely use public transport, 61% shift to pattern 2 which has a higher travel frequency.

Peak-hour Activity

Among the affected passengers, those who did not often travel during peak hours (pattern 4) are more likely to shift to patterns with higher activity during peak hours (patterns 1, 2, and 3) compared to the reference group. One possible reason for this is that the disruptions in our case study primarily occurred during off-peak hours. As a result, some passengers might adjust their travel timing to avoid potential disruptions similar to what they experienced during off-peak hours. Specifically, 62% of the affected passengers in pattern 4 (off-peak travelers) transition to the first three groups (peak travelers), whereas this value is only 30% for the reference group.

Travel Flexibility

We observe that, in both the affected and reference groups, patterns with less flexibility in first and last boarding times (patterns 1 and 2) exhibit higher travel behavior inertia compared to more flexible patterns (patterns 3 and 4). One possible reason for this could be that the habitual behavior of traveling at the same time every day has extended to their overall travel patterns, making them more likely to stick with familiar routines. As a result, they tend to exhibit higher inertia and are less likely to change their travel behavior.

Overall Comparison of Affected and Reference Groups

In conclusion, the differences that exist between the reference and affected passengers are not very substantial, whereas the extent of this difference was expected to be more pronounced given the occurrence of the disruption. It is evident that this difference is less substantial for the first three patterns, which comprise 80% of the data. The main differences are observed in patterns 4 (Occasional very flexible travellers (weekend)) and pattern 5 (Sporadic travellers) which are the smaller patterns. It is possible that if the disruptions were more significant, the differences would become more pronounced.

5.2.3. Analysis of the Mobility Styles

The differences in the temporal dynamics of individuals' travel patterns (states) are captured by the mobility styles (classes). In other words, passengers are grouped into various mobility styles based on the differences and similarities in their travel pattern transition probabilities over time. Unlike travel patterns, which can change for a passenger over time, these mobility styles are constant over time.

These different mobility styles within the population can be identified using the mixture latent Markov model. Unlike simple latent Markov models, which assume all passengers belong to a single mobility style and thus share homogeneous transition probabilities, mixture latent Markov models account for heterogeneity by assuming the population consists of multiple latent mobility styles. Each style follows its own Markov process with unique transition probabilities. For example, a simple Markov model would not differentiate between a group of passengers that decreases their travel frequency over time and another group that maintains their frequency. In contrast, the mixture model distinguishes between these groups, revealing their differing tendencies.

This model is particularly useful in our study because it allows us to test specific expectations. For example, we anticipate that affected passengers will be more or less likely to belong to certain mobility styles, or that some mobility styles will emerge while others may not form due to the disruption. The mixture latent Markov model enables us to test these hypotheses by identifying the different mobility styles, analyzing their characteristics and sizes, and drawing conclusions about the behavioral changes resulting from the disruption.

Description of the Mobility Styles

In this section the purpose is to use the transition probability matrices of the affected and the reference passengers (Tables 5.5, 5.6, and 5.7), to provide information on each mobility style. This helps us in identifying the different behavior traits that exist among the passengers. Later in Section 5.2.3 and 6 the definition of the mobility styles is used to draw conclusions about the differences among the reference and the affected passengers.

Mobility style 1 (stayers)

The first mobility style is labeled as ”stayers” because passengers in this class maintain the same travel pattern across all periods and never change their behavior. This is evident from Table Table 5.5, which shows a likelihood of 1 for staying in the same pattern and 0 for transitioning to other patterns.

Table 5.5: Transition matrices of mobility style 1.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**1**	0	0	0	0
	2. Occasional less flexible peak travellers	0	**1**	0	0	0
	3. Frequent flexible peak travellers (weekend)	0	0	**1**	0	0
	4. Occasional very flexible travellers (weekend)	0	0	0	**1**	0
	5. Sporadic travellers	0	0	0	0	**1**
Reference	1. Frequent less flexible peak travellers	**1**	0	0	0	0
	2. Occasional less flexible peak travellers	0	**1**	0	0	0
	3. Frequent flexible peak travellers (weekend)	0	0	**1**	0	0
	4. Occasional very flexible travellers (weekend)	0	0	0	**1**	0
	5. Sporadic travellers	0	0	0	0	**1**

Mobility style 2

This class is inclined towards travelling more frequently, less flexibly, more during peak, and less in the weekend. This inclination is more pronounced among the affected passengers than the reference:

- Higher frequency: Passengers tend to shift towards traveling more frequently. Based on Table 5.6, frequent patterns (1 and 3), in both the affected and reference groups, are most likely to transition within the frequent patterns 1 and 3. This is due to the large transition probabilities of pattern 1 to 1 (80% among the affected and 67% among the reference), as well as pattern 3 to pattern 1 (82% among the affected and 63% among the reference) and pattern 3 to 3 (16% among the affected and 21% among the reference). Moreover, in the affected group, pattern 2 (the largest occasional pattern) is more inclined towards transitioning to the frequent pattern 1 rather than staying in the same pattern. Moreover, in the affected group, pattern 5 (sporadic travelers) shifts to pattern 2 (occasional travelers with higher frequency) with a 78% probability. In the reference group, however, the behavior of the occasional and sporadic patterns is more inclined towards maintaining their frequency.

- Less flexibility: Passengers tend to become less flexible regarding boarding times. In both the affected and reference groups, the sum of the probabilities of shifting to less flexible patterns (1 and 2) is significantly higher than shifting to more flexible patterns (3 and 4). This trend is less pronounced among the reference group.
- More peak-hour activity: Passengers are more likely to transition to patterns with higher activity during peak hours. This is evident from the large transition probabilities towards patterns 1, 2, and 3, which have high activity during peak hours, with pattern 1 being the most active pattern during peak.
- Less weekend activity: Passengers are more likely to transition to patterns with less weekend activity. This is shown by the large transition probabilities towards patterns 1 and 2, which have no activity on weekends.

Table 5.6: Transition matrices of mobility style 2.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.798**	0.124	0.077	0.001	0.000
	2. Occasional less flexible peak travellers	**0.490**	**0.426**	0.008	0.032	0.043
	3. Frequent flexible peak travellers (weekend)	**0.818**	0.011	**0.162**	0.004	0.006
	4. Occasional very flexible travellers (weekend)	**0.168**	**0.807**	0.001	0.023	0.002
	5. Sporadic travellers	0.052	**0.779**	0.002	0.002	**0.166**
Reference	1. Frequent less flexible peak travellers	**0.668**	0.203	0.108	0.006	0.015
	2. Occasional less flexible peak travellers	**0.161**	**0.649**	0.020	0.059	0.111
	3. Frequent flexible peak travellers (weekend)	**0.629**	0.134	**0.215**	0.011	0.011
	4. Occasional very flexible travellers (weekend)	0.014	**0.772**	0.001	0.070	**0.144**
	5. Sporadic travellers	0.035	**0.366**	0.006	0.057	**0.537**

Mobility style 3

This class is inclined towards travelling more flexibly, less during peak, and more in the weekend while travelling with the same frequency. This inclination is more pronounced among the affected than the reference:

- Same frequency: Passengers tend to maintain their travel frequency. We can see that the frequent patterns (1 and 3) are more likely to transition within the frequent patterns (1 and 3). For pattern 1, the sum of transition probabilities to patterns 1 and 3 is 94% among the affected, and 84% among the reference. For pattern 3, the sum of transition probabilities to patterns 1 and 3 is 86% among the affected, and 65% among the reference. Similarly, the occasional patterns (2 and 4) majorly transition to occasional patterns (2 and 4). The sporadic travellers (pattern 5) either remain as sporadic travellers or transition to pattern 2 and 4 (occasional travellers) and thus increase their frequency slightly.
- More flexibility: Passengers tend to become more flexible regarding the boarding times. For example among the affected and the reference groups, passengers in pattern 1 (less flexible) are most likely to transition to pattern 3 (more flexible) rather than staying in the same pattern which is less flexible. The probability of the transition from pattern 1 to 3 is 58% among the affected and 61% among the reference group. Similarly, pattern 2 (less flexible) transitions to pattern 4 (more flexible) with a probability of 39% among the affected and 41% among the reference.

- Less peak-hour activity: Passengers tend to travel less often during peak. For example pattern 1 (most active pattern during peak) most likely transitions to pattern 3 which has less peak activity compared to pattern 1 and 2. Pattern 2 most often transitions to pattern 4 which has even lower activity in peak. Finally, pattern 3 mostly remains as is.
- More weekend activity: Passengers tend to travel more often during the weekend. We observe substantial transitions from patterns 1 and 2 (with no weekend activity) to patterns 3 and 4 which are active in the weekend.

Table 5.7: Transition matrices of mobility style 3.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.359**	0.017	**0.582**	0.042	0.000
	2. Occasional less flexible peak travellers	0.041	**0.398**	0.065	**0.392**	0.104
	3. Frequent flexible peak travellers (weekend)	**0.164**	0.013	**0.690**	0.109	0.024
	4. Occasional very flexible travellers (weekend)	0.116	**0.373**	0.173	**0.335**	0.003
	5. Sporadic travellers	0.000	**0.727**	0.016	0.016	**0.241**
Reference	1. Frequent less flexible peak travellers	**0.225**	0.021	**0.611**	0.142	0.000
	2. Occasional less flexible peak travellers	0.008	**0.344**	0.088	**0.407**	**0.152**
	3. Frequent flexible peak travellers (weekend)	0.079	0.102	**0.573**	**0.217**	0.029
	4. Occasional very flexible travellers (weekend)	0.006	**0.209**	0.072	**0.599**	0.114
	5. Sporadic travellers	0.000	**0.191**	0.036	**0.337**	**0.436**

Discussion of the Mobility Styles and Initial Travel Patterns Sizes

The probability of mobility style membership, i.e., the total size, is presented in Table 5.8 for the affected and reference groups. Comparing these probabilities among the affected and reference passengers, we do not observe substantial differences which is against our expectations. First, we expected that the affected passengers would be less likely to belong to the stayers class (passengers maintaining the same travel pattern) compared to the reference group. Our expectation was based on the assumption that passengers who experienced the disruption would be more likely to alter their behavior, while those in the reference group, who did not experience any disruption, would be more inclined to maintain their existing travel patterns. However, the likelihood of being a stayer is 27% among the affected group and 24% among the reference, indicating that the likelihood of never changing behavior is quite similar for both groups.

Furthermore, we anticipated observing a mobility style characterized by a tendency toward traveling less frequently, and we expected that affected passengers would be more likely to belong to this class compared to the reference group. However, this expected mobility style did not emerge. Instead, mobility style 2 and 3 either maintain the same travel frequency or increase it.

Moreover, the affected passengers are just as likely (46%) to belong to mobility style 2 (tendency towards traveling more frequently) as the reference group (46%). This indicates that the tendency towards traveling more frequently does not differ between the affected passengers and the reference group. This finding contradicts our expectation that the affected passengers would be more likely to travel less frequently. We expected that this value would be significantly lower among the affected passengers compared to the reference.

Regarding the third mobility style, which consists of passengers with a tendency towards off-peak travel, the class sizes are again quite similar. However, the 3% difference between the affected and the reference passengers might be due to the fact that the disruptions in this case study often occur during off-peak hours. Therefore, the affected passengers are slightly less inclined towards traveling during off-peak hours, as they experienced multiple disruptions during those times.

Table 5.8: Mobility style size of the affected and reference passengers.

	Total size	
Mobility style	affected	reference
1 (stayer)	0.267	0.238
2	0.462	0.456
3	0.272	0.307

Next, the initial travel pattern sizes (in the pre-disruption period (t = 0)) are presented in Table 5.9 for both the affected and the reference passengers. Based on these initial travel pattern sizes for the affected passengers in the first row, it is evident that pattern 1 (frequent less flexible peak travellers) is most strongly represented in the sample (59%). The next largest travel pattern is the occasional less flexible peak travellers (19%) followed by pattern 3, the frequent flexible peak travellers (17%), and the occasional very flexible travellers (5%). The smallest pattern is the sporadic travellers (0.1%).

This distribution differs somewhat among the reference passengers, as seen in Table 5.9. Pattern 1 is not as strongly represented among the reference group as it is among the affected group. Instead, patterns 2, 3, and 4 are more strongly represented in the reference group compared to the affected group. Ideally, the reference group would have a similar initial distribution of travel patterns as the affected passengers, making the reference group more representative and the comparison more reliable. However, some differences between the two groups are inevitable since the reference passengers are different individuals from the affected passengers. One potential reason for this difference is the process used to identify the reference and affected passengers. Reference passengers are those who have no overlapping journeys with the disruptions, and depending on the day, time, and number of disruptions, this can filter out certain groups of passengers, leading to differences in the composition of the two groups.

Table 5.9: Travel pattern size of the affected and reference passengers at t = 0 (pre-disruption).

Total state size (t = 0)	1. Frequent less flexible peak travellers	2. Occasional less flexible peak travellers	3. Frequent flexible peak travellers (with weekend)	4. Occasional very flexible travellers (with weekend)	5. Sporadic travellers
Affected	0.594	0.186	0.171	0.048	0.001
Reference	0.380	0.215	0.281	0.124	0.000

6
Discussion and Conclusion

In this study, we proposed a framework to investigate the long-term effects of public transport disruptions on passenger behavior and applied it to a case study in the Washington D.C. metro network. Using a mixture latent Markov model, we identified five distinct travel patterns and monitored the transitions of passengers between these patterns over time. Based on these travel patterns and transition probabilities, we categorized passengers into three mobility styles: (1) those who maintain the same travel behavior over time, (2) those who tend to travel more frequently, less flexibly, more often during peak hours, and less on weekends, and (3) those who travel more flexibly, less during peak hours, and more on weekends, while maintaining the same frequency of travel.

The analysis of the results reveals that the difference between the affected and reference passengers is not as substantial as initially expected, especially regarding the three largest travel patterns (patterns 1, 2, and 3). Some differences are observed in pattern 4, the occasional very flexible travelers, and pattern 5, the sporadic travelers, but these are the smallest patterns and consist of only a small proportion of the passengers.

The insignificance of the difference between the affected and reference passengers could indicate that the disruption studied did not have a substantial effect on passenger behavior. This might be because the disruptions had a low frequency and duration, leading to only temporary delays and therefore having little long-term impact. This observation aligns with previous findings in the literature. For example, Drabicki et al. (2021) found that the higher the frequency of disruptions, the more likely it is to result in long-term behavioral changes among passengers. In their study, 77% reported changing their route or departure time to avoid routes/times that are frequently disrupted. Moreover, in a study by Papangelis et al. (2016), passengers also reported making permanent behavioral changes, such as mode changes, relocation, and job changes due to experiencing frequent long-term disruptions in the past. Tian and Zheng (2018) also found that in response to minor delays (of less than 30 minutes) leading to frequency reductions, the majority (86%) of passengers remained in the system and either waited or postponed their departure time (if not in the system when the disruption occurred).

Another reason for this minimal impact could be that the disruptions in our case study primarily occur during off-peak hours. According to the literature, one of the determining factors in passengers' responses to disruptions is whether the disruption happens during peak or off-peak hours. If disruptions occur during peak hours, passengers are more likely to change their travel mode and opt for taxis, buses, or bikes instead of waiting (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018). This behavior is often driven by the need to arrive at their destinations, such as work or school, on time. Li, Yao, Yamamoto, Tang, and Liu (2020) also found that during peak hours, passengers perceive travel time more negatively and are willing to pay more to reduce it. Although these observations are based on passengers' immediate responses to disruptions, they might extend to the long term if disruptions are repeated and frequent, as observed by Drabicki et al. (2021) and Papangelis et al. (2016).

Based on the transition probabilities in Table 5.4, we observed a high level of travel behavior inertia, i.e., the tendency to maintain the same travel pattern over time, among the affected passengers. This inertia is most pronounced in the three largest travel patterns (patterns 1, 2, and 3), where passengers remain in the same travel pattern between consecutive periods with a probability of at least 67%. The largest pattern, pattern 1 (frequent less flexible peak travelers), exhibits the highest inertia at 85%.

We initially expected that passengers experiencing disruptions would be more likely to adjust and change their behavior, resulting in lower travel behavior inertia compared to those who do not experience disruptions. However, Table 5.4 reveals that while this expectation holds for most patterns (patterns 2, 4, and 5), where the transition probability of staying in the same travel pattern is lower among affected passengers compared to the reference group, it does not hold for patterns 1 and 3.

We also anticipated that frequent travelers affected by the disruptions would transition to occasional travel patterns and thus travel less frequently due to the inconvenience they experienced because of the disruptions. For instance, they might choose to work from home on some days or drive their cars to reach their destinations. However, our findings did not support this expectation. Instead, the affected frequent travelers (patterns 1 and 3) exhibited high travel behavior inertia, continuing to travel as frequently as before, as evident from Table 5.4. This observation aligns with the findings of Adelé et al. (2019) and Li, Yao, Yamamoto, Tang, and Liu (2020), who noted that frequent travelers are less likely to change their routes or modes when facing disruptions and are more likely to wait for the service to resume.

Several factors could explain this behavior. Frequent travelers might be employees who are required to commute to the office daily without the option of working from home, or they may not own a car, limiting their alternatives to public transport. As a result, they are unable to reduce their reliance on public transport even when disruptions occur. Additionally, their frequent use of public transport might have become habitual, making them less sensitive to disruptions. Adelé et al. (2019) suggests that, according to the theory of bounded rationality, individuals tend to seek solutions that are ”good enough” rather than optimal. They simplify their decision-making processes by balancing the time and effort required to find a solution with the perceived quality of that solution. This may explain why frequent travelers tend to stick with their usual behaviors, despite disruptions.

In contrast, we observe less travel behavior inertia among occasional and sporadic travelers (patterns 2, 4, and 5) compared to frequent travelers (patterns 1 and 3). This may be because occasional travelers have more flexibility, such as the option to work from home or access to personal vehicles, which might be why they initially use public transportation less frequently. Furthermore, their use of public transport may be less habitual, making them more responsive to disruptions.

Regarding the sizes of mobility styles presented in Table 5.8, we did not observe substantial differences between the affected and the reference groups, which contradicts our expectations. Specifically, we anticipated a smaller size for mobility style 1 (the stayers) among the affected passengers compared to the reference group. This expectation was based on the assumption that affected passengers, having experienced disruptions, would be more likely to modify their behavior and thus be less likely to maintain the same travel pattern across periods. In contrast, we expected reference passengers, who did not experience disruptions, to have a higher likelihood of staying in the same travel pattern.

Furthermore, both the affected and reference groups show similar likelihoods of belonging to mobility style 2 (tendency towards traveling more frequently and less flexibly). We expected this likelihood to be smaller among the affected passengers, as we hypothesized that they would travel less frequently due to the inconveniences caused by disruptions.

The likelihood of membership in the third mobility style, which consists of passengers with a tendency towards off-peak travel, is also quite similar between the reference and affected groups. The slightly smaller value of 3% among the affected passengers could be attributed to the fact that our disruptions often occur during off-peak hours. Consequently, the affected passengers are slightly less inclined to travel during off-peak hours, making them less likely to be members of mobility style 3. Lastly, the absence of a mobility style characterized by a tendency towards less frequent travel might be another indication of the minimal impact of the disruptions observed in our case study.

Lastly, we observed that the model, which included a dummy variable as a covariate to differentiate between the transition probabilities from the pre-disruption period to the first post-disruption period and those from other post-disruption periods, did not converge to an optimal solution even after 200 iterations and 200 different parameter starting sets. This indicates that the model could not identify a proper set of parameters to define the relationships effectively. As a result, we relaxed the constraint of varying transition probabilities and instead assumed that these probabilities are consistent across all pairs of consecutive periods. One possible explanation for this outcome is that the disruption may not have had a significant effect, leading to minimal differences in transitions between pairs of periods.

6.1. Limitations and Future Research

One of the limitations of this study was the availability of detailed data on the disruptions, specifically accurate data regarding the start and end times of the disruptions. Due to this limitation, we had to infer the occurrence of disruptions based on train movements from the AVL dataset, which required extensive data cleaning and special handling.

Another limitation was the restricted effective period available for analyzing disruptions. The onset of COVID-19 in Washington, DC, in March 2020 limited our analysis to a seven-month period before the pandemic. Ideally, we would prioritize analyzing significant disruptions, such as station or line closures, over frequency reductions, as these are more likely to have a substantial long-term impact. However, such significant disruptions were not present in the available data for the case study period.

Additionally, the unavailability of socio-demographic data was a limitation. Understanding the impact of factors such as income, age, education, and car ownership on passengers' long-term travel behavior would be valuable. Literature suggests that these characteristics significantly influence passengers' immediate responses to disruptions. Therefore, evaluating their impact in the long term would also be worthwhile.

Future research can extend the current methodology by analyzing changes in the spatial aspect of travel behavior in addition to the temporal aspect. This is important because some passengers might maintain their usual travel times but alter their travel routes. For instance, regular commuters may not be able to change their departure times or reduce their travel frequency due to the need to be at work at the same time every day. However, they might adjust their routes by traveling from different stations than usual. Researchers can explore whether passengers change their frequent boarding and alighting stations after a significant disruption or if they use these stations less frequently. For example, if a regular station is disrupted for an extended period, passengers might switch to nearby, undisrupted stations instead. It would be interesting to determine whether passengers return to their initial stations once the disruption is resolved, to what extent their usage levels recover, or if they continue using the new stations.

A second avenue for future research is incorporating socio-demographic information, such as age, income, education, and car ownership as covariates in the model to evaluate the impact of these characteristics on transition probabilities, initial state membership, and class membership. Understanding the socio-demographic composition of each class can provide insights into why certain travel patterns emerge, such as preferences for peak versus off-peak travel. For instance, younger individuals might prefer off-peak hours due to flexible schedules, while older adults might travel during specific times for routine activities. These insights can inform targeted strategies to optimize transport services and address the specific needs of different user groups.

Moreover, other references could be explored to determine whether they would more accurately represent the affected passengers' behavior under normal conditions. One potential reference could be the affected passengers' own behavior during the same period in the previous year. For example, if the pre-disruption period is in January 2024, the behavior of affected passengers in January 2023 could be used as the reference for the pre-disruption period.

Lastly, the framework can be applied to other case studies, particularly those involving more significant disruptions such as long-term station or line closures. Additionally, the impact of the characteristics of the disruption can be studied by replicating the methodology for different disruptions that vary in magnitude and duration. For example, results could indicate that when station or line closures last longer than a certain number of months, the changes in passenger travel behavior are less likely to be reversible. This could provide valuable insights for public transportation operators when planning major maintenance works.

6.2. Recommendations

There are several recommendations to make the analysis more comprehensive and reliable. One key recommendation is to collect detailed data about the disruptions, including start time, end time, location, and impact, and use this information instead of inferring disruption times from the AVL data to identify affected passengers. Relying on AVL data to infer disruptions, as was done in this study, might negatively influence the results if the AVL data quality is insufficient. For instance, if a system malfunction fails to record train movements, the period might be mistakenly counted as a disruption because no trains are observed serving a station. Additionally, AVL data often require special handling and preparation, such as addressing missing values in some columns, as was necessary in this study.

Another recommendation, if disruption data is available, is to incorporate the cause of the disruption as a factor when analyzing passenger behavior changes. For example, examining whether disruptions caused by system failures have different impacts compared to those caused by severe weather could provide valuable insights. Passengers might be more annoyed by disruptions resulting from system failures and malfunctions because they may perceive these issues as a sign of poor maintenance by PT operators, potentially leading to a loss of trust in the reliability of public transportation. In contrast, disruptions caused by accidents, severe weather, or other factors beyond the control of PT operators might result in a lesser loss of trust, as passengers might be more understanding of such uncontrollable events.

Lastly, when other studies implement our framework, practical recommendations can be derived from the results. For example, if the results show that a group of passengers reduces their use of public transport after a major disruption, this observation can provide evidence for public transport authorities that they would need to not only implement mitigating measures immediately after a disruption, but also extend these measures in the long-term to prevent passengers from abandoning the system. For instance, authorities could offer discounts to passengers who were highly affected by the disruptions, especially those in suburban areas with limited alternative modes of transport who might have experienced greater frustration due to the disruptions. The goal would be to provide an incentive for these passengers to continue using public transport as frequently.

6.3. Conclusion

In conclusion, this study aimed to develop a framework for gaining a deep understanding of the prolonged effects of public transport disruptions on passengers' travel behavior by extending the analysis to five to six months post-disruption, in contrast to most previous studies that focus only on immediate impacts. To achieve this, we applied a mixture latent Markov model (MLMM) to uncover the unobserved travel patterns of passengers across six periods. This analysis was based on a range of indicators, including (1) share of active days, (2) average number of journeys per weekday, (3) share of weekend journeys, (4) share of peak-hour journeys, (5) share of similar first boarding times, and (6) share of similar last boarding times. The selection of these indicators was informed by our expectations regarding potential passenger responses to disruptions, a review of relevant literature, and data availability.

Using these indicators five travel patterns were identified that captured travel behavior in each period:

- *Frequent less flexible peak travellers:* These passengers travel frequently with high peak-hour activity and moderate flexibility in their boarding times.
- *Occasional less flexible peak travellers:* These passengers travel occasionally, primarily during peak hours, with moderate flexibility in their boarding times.
- *Frequent flexible peak travellers (with weekend):* These passengers are the most active, traveling frequently on both weekdays and weekends, with peak-hour activity and relatively flexible boarding times.
- *Occasional very flexible travellers (with weekend):* These passengers travel occasionally with significant weekend activity, less peak-hour activity, and the highest flexibility in their boarding times.
- *Sporadic travellers:* These passengers rarely travel, with minimal activity on both weekdays and weekends, and consistent boarding times.

Given that the MLMM accounts for heterogeneity in transition probabilities among individuals—unlike standard latent Markov models which assume homogeneity—we specifically employed the MLMM to uncover distinct classes, or mobility styles, among the passengers. As a result, the MLMM identified three mobility styles: (1) those who maintain the same travel behavior over time, (2) those with a tendency towards traveling more frequently, less flexibly, more during peak, and less in the weekend, and (3) those with a tendency towards travelling more flexibly, less during peak, and more in the weekend while travelling at the same frequency level.

The analysis of transition probabilities and mobility styles reveals that the long-term effects of the case study disruption on passenger travel behavior were less significant than initially expected. The observed differences between affected and reference passengers in terms of transition probabilities and mobility style membership were minimal and did not align with our initial expectations. Notably, a high level of travel behavior inertia was observed among the affected passengers, contrary to our anticipation that they would exhibit lower inertia due to their experience of disruption, which was expected to prompt modifications in their behavior. This suggests that the disruptions had a limited impact on altering the travel behaviors of these passengers. Additionally, we observed that inertia was highest among frequent travelers, an observation that is consistent with current literature. Factors such as the low frequency and duration of the disruptions, their occurrence during off-peak hours, and the habitual nature of travelers' routines likely contributed to this outcome.

These findings underscore the complexity of travel behavior adaptation and highlight the need for further research on the interplay between disruption characteristics and passenger responses. The framework however can successfully reveal the travel patterns and the mobility styles and thus can be easily applied to other case studies.

References

Adelé, S., Tréfond-Alexandre, S., Dionisio, C., & Hoyau, P.-A. (2019). Exploring the behavior of suburban train users in the event of disruptions. *Transportation research part F: traffic psychology and behaviour*, *65*, 344–362.

Arslan Asim, M., Weiss, A., Kattan, L., & Wirasinghe, S. (2021). Transit users' mode choice behavior during light rail transit short-term planned service disruption. *Transportation Research Record*, *2675*(10), 711–722.

Bartolucci, F., Farcomeni, A., & Pennoni, F. (2010). An overview of latent markov models for longitudinal categorical data. *arXiv preprint arXiv:1003.2804*.

Bhaskar, A., Chung, E., et al. (2014). Passenger segmentation using smart card data. *IEEE Transactions on intelligent transportation systems*, *16*(3), 1537–1548.

Briand, A.-S., Côme, E., Trépanier, M., & Oukhellou, L. (2017). Analyzing year-to-year changes in public transport passenger behaviour using smart card data. *Transportation Research Part C: Emerging Technologies*, *79*, 274–289.

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

Eltved, M., Breyer, N., Ingvardson, J. B., & Nielsen, O. A. (2021). Impacts of long-term service disruptions on passenger travel behaviour: A smart card analysis from the greater copenhagen area. *Transportation Research Part C: Emerging Technologies*, *131*, 103198.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd*, *96*(34), 226–231.

Ghaemi, M. S., Agard, B., Trépanier, M., & Partovi Nia, V. (2017). A visual segmentation method for temporal smart card data. *Transportmetrica A: Transport Science*, *13*(5), 381–404.

Gudicha, D. W., Schmittmann, V. D., & Vermunt, J. K. (2017). Statistical power of likelihood ratio and wald tests in latent class models with covariates. *Behavior research methods*, *49*, 1824–1837.

Haustein, S., & Kroesen, M. (2022). Shifting to more sustainable mobility styles: A latent transition approach. *Journal of Transport Geography*, *103*, 103394.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178–210.

Karlaftis, M. G., Kepaptsoglou, K., Stathopoulos, A., & Dimitriou, D. (2006). *Public transportation during the athens 2004 olympics: From planning to performance evaluation* (tech. rep.).

Kroesen, M., & van Cranenburgh, S. (2016). Revealing transition patterns between mono-and multimodal travel patterns over time: A mover-stayer model. *European Journal of Transport and Infrastructure Research*, *16*(4), 754–771.

Lathia, N., Smith, C., Froehlich, J., & Capra, L. (2013). Individuals among commuters: Building personalised transport information services from fare collection systems. *Pervasive and Mobile Computing*, *9*(5), 643–664.

Li, B., Yao, E., Yamamoto, T., Huan, N., & Liu, S. (2020). Passenger travel behavior analysis under unplanned metro service disruption: Using stated preference data in guangzhou, china. *Journal of Transportation Engineering, Part A: Systems*, *146*(2), 04019069.

Li, B., Yao, E., Yamamoto, T., Tang, Y., & Liu, S. (2020). Exploring behavioral heterogeneities of metro passenger's travel plan choice under unplanned service disruption with uncertainty. *Transportation Research Part A: Policy and Practice*, *141*, 294–306.

Liu, T., Ma, Z., & Koutsopoulos, H. N. (2021). Unplanned disruption analysis in urban railway systems using smart card data. *Urban Rail Transit*, *7*(3), 177–190.

Ma, X., Liu, C., Wen, H., Wang, Y., & Wu, Y.-J. (2017). Understanding commuting patterns using transit smart card data. *Journal of Transport Geography*, *58*, 135–145.

Ma, X., Wu, Y.-J., Wang, Y., Chen, F., & Liu, J. (2013). Mining smart card data for transit riders' travel patterns. *Transportation Research Part C: Emerging Technologies*, *36*, 1–12.

Magidson, J., & Vermunt, J. (2002). Latent class models for clustering: A comparison with k-means. *Canadian journal of marketing research*, *20*(1), 36–43.
Magidson, J., & Vermunt, J. K. (2004). Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, 175–198.
Magidson, J., Vermunt, J. K., & Tran, B. (2009). Using a mixture latent markov model to analyze longitudinal us employment data involving measurement error. *New trends in psychometrics*, 235–242.
Medina, S. A. O. (2018). Inferring weekly primary activity patterns using public transport smart card data and a household travel survey. *Travel Behaviour and Society*, *12*, 93–101.
Mo, B., von Franque, M. Y., Koutsopoulosc, H. N., Attanuccid, J., & Zhao, J. (2022). Impact of unplanned service disruptions on urban public transit systems. *arXiv preprint arXiv:2201.01229*.
Murray-Tuite, P., Wernstedt, K., & Yin, W. (2014). Behavioral shifts after a fatal rapid transit accident: A multinomial logit model. *Transportation research part F: traffic psychology and behaviour*, *24*, 218–230.
Nazem, M., Lomone, A., Chu, A., & Spurr, T. (2018). Analysis of travel pattern changes due to a medium-term disruption on public transit networks using smart card data. *Transportation Research Procedia*, *32*, 585–596.
Nguyen-Phuoc, D. Q., Currie, G., De Gruyter, C., & Young, W. (2018a). How do public transport users adjust their travel behaviour if public transport ceases? a qualitative study. *Transportation research part F: traffic psychology and behaviour*, *54*, 1–14.
Nguyen-Phuoc, D. Q., Currie, G., De Gruyter, C., & Young, W. (2018b). Transit user reactions to major service withdrawal–a behavioural study. *Transport Policy*, *64*, 29–37.
Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural equation modeling: A multidisciplinary Journal*, *14*(4), 535–569.
Ou, Y., Cai, C., et al. (2018). Large-scale transit market segmentation with spatial-behavioural features. *Transportation Research Part C: Emerging Technologies*, *90*, 97–113.
Papangelis, K., Velaga, N. R., Ashmore, F., Sripada, S., Nelson, J. D., & Beecroft, M. (2016). Exploring the rural passenger experience, information needs and decision making during public transport disruption. *Research in Transportation Business & Management*, *18*, 57–69.
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), 258–271.
Ran, X., Xi, Y., Lu, Y., Wang, X., & Lu, Z. (2023). Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*, *56*(8), 8219–8264.
Sander, J., Ester, M., Kriegel, H.-P., & Xu, X. (1998). Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data mining and knowledge discovery*, *2*, 169–194.
Saxena, N., Rashidi, T. H., & Auld, J. (2019). Studying the tastes effecting mode choice behavior of travelers under transit service disruptions. *Travel behaviour and society*, *17*, 86–95.
Schiphol. (2022). *Third place in europe based on total air transport movements* [Accessed on March 06, 2024]. https://www.schipholannualtrafficreview.nl/2022/other-airports
Shires, J. D., Ojeda-Cabral, M., & Wardman, M. (2019). The impact of planned disruptions on rail passenger demand. *Transportation*, *46*(5), 1807–1837.
Tian, X., & Zheng, B. (2018). Using smart card data to model commuters' responses upon unexpected train delays. *2018 IEEE International Conference on Big Data (Big Data)*, 831–840.
Vermunt, J. K., & Magidson, J. (2013). Technical guide for latent gold 5.0: Basic, advanced, and syntax. *Belmont, MA: Statistical Innovations Inc*.
Wang, L., Wang, Y., Sun, X., Wu, Y., Peng, F., Chen, C.-H. P., & Song, G. (2023). Public transit passenger profiling by using large-scale smart card data. *Journal of Transportation Engineering, Part A: Systems*, *149*(4), 04023013.
Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: A guide to best practice. *Journal of Black Psychology*, *46*(4), 287–311.
WMATA. (2023). *More service coming to the red line during weekday rush hours starting monday, may 8* [Accessed on April 7, 2024]. https://www.wmata.com/about/news/Red-Line-May-2023-Improvements.cfm

WMATA. (2024a). *Metro ridership snapshot: March 2020* [Accessed on March 21, 2024]. https://www.wmata.com/initiatives/ridership-portal/upload/March-2020-Ridership-Snapshot.pdf

WMATA. (2024b). *Metrorail* [Accessed on February 9, 2024]. https://www.wmata.com/service/rail/

WMATA. (2024c). *Milestones history* [Accessed on February 9, 2024]. https://www.wmata.com/about/history/index.cfm

WMATA. (2024d). *Timetables: Rail schedules* [Accessed on May 20, 2024]. https://www.wmata.com/schedules/timetables/

Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of data science*, *2*, 165–193.

Zhao, J., Qu, Q., Zhang, F., Xu, C., & Liu, S. (2017). Spatio-temporal analysis of passenger travel patterns in massive smart card data. *IEEE Transactions on Intelligent Transportation Systems*, *18*(11), 3135–3146.

Zhu, S., Masud, H., Xiong, C., Yang, Z., Pan, Y., & Zhang, L. (2017). Travel behavior reactions to transit service disruptions: Study of metro safetrack projects in washington, dc. *Transportation Research Record*, *2649*(1), 79–88.

A

Scientific paper

Investigating the Long-Term Impact of Disruptions on Passenger Travel Behavior Using AFC Data

A Case Study of Washington D.C. Metro Network

Fariba Tavakoli[1], Oded Cats[1], Yongqiu Zhu[1], Maarten Kroesen[2]

[1] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, the Netherlands
[2] Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands

Abstract

Disruptions in urban public transport networks significantly impact urban mobility, causing inconveniences for passengers, economic losses, and vulnerabilities in transit systems. Major disruptions may lead some passengers to permanently switch to alternative modes like cars, even after the disruptions resolve, potentially worsening congestion and environmental issues. While the immediate impact of disruptions on behavior has been extensively studied, there is limited research on the long-term effects, often restricted to qualitative methods. This study aims to address this gap by proposing a framework to investigate the prolonged effects of public transport disruptions on passenger travel behavior using smart card data. A mixture latent Markov model is used to track passenger behavior from the pre-disruption to the post-disruption period. This framework identifies travel patterns and tracks how passengers transition between these patterns over time, thus inferring the impact of disruptions on behavior. Each passenger is assigned a mobility style that reflects their general travel attitude, such as those who do not change their behavior. The results from our case study reveal that the impact of the disruption was not as substantial as anticipated, with a high proportion of passengers maintaining their behavior.

Keywords: Disruption, public transport, long-term effect, travel behavior, AFC, mixture latent Markov

1. Introduction

In urban public transport networks, disruptions are inevitable, often arising from maintenance works, signal failures, weather conditions, and other causes. These disruptions significantly impact urban mobility, leading to economic losses and increased vulnerabilities within transit networks. They also affect passenger travel behavior, influencing mode and route choices, departure times, and causing trip cancellations. Major consecutive disruptions may result in passengers who experimented with alternative travel modes, such as cars, during disruptions to permanently switch to these new modes even after the disruptions resolve (Karlaftis et al., 2006) (Zhu et al., 2017). The rise of sharing-economy travel options, such as ride-hailing and car sharing, further intensifies competition with public transit, encouraging more people to opt for these alternatives (Rahimi et al., 2020). If passengers shift towards cars due to the resulting inconvenience, it can exacerbate congestion and environmental issues.

1.1. Scientific Gap and Contributions

This study fills a gap in the literature by introducing a framework to investigate long-term behavioral changes, specifically extending five to six months after a disruption. While immediate changes in passenger behavior are extensively studied, long-term impacts remain underexplored. Understanding these long-term impacts is crucial because it reveals whether temporary changes in behavior become permanent. Additionally, it allows for an examination of passengers' tolerance towards disruptions and the extent to which they maintain their travel behavior, demonstrating travel behavior inertia after a disruption. If results indicate that passengers leave the system or reduce their usage after a disruption, the findings can guide the need for extended mitigation measures beyond the immediate aftermath to prevent this decline in usage.

The aim of this study is to develop a general framework for understanding how passengers change their behavior in the long term after a disruption. To achieve this, we cluster passengers' travel behavior across different periods to uncover the unobserved travel patterns. By examining how passengers transition between these travel patterns, we analyze the impact of the disruption. One of our main contributions is that we use a mixture latent Markov model, which is particularly suitable for clustering longitudinal data and probabilistically assigns data to clusters, accounting for uncertainties in the process.

The remainder of this paper is organized as follows: Section 2 discusses previous work on the impact of disruptions on passenger behavior and the influencing factors. Section 3 explains the methodology and section 4 provides information on the case study and the results. Finally, Section 5 provides a discussion of the results and directions for future research.

2. Related Literature

Most studies on the impact of transport disruptions focus on immediate passenger responses, with few examining long-term behavior. Surveys are the most common data source, while studies using AFC transaction data are rare. Research by Nazem et al. (2018) and Eltved et al. (2021), using AFC data, indicates that station ridership and travel frequency do not return to pre-disruption levels even months after the disruption ends. However, these studies need to extend the post-disruption period to determine if the impacts persist long-term and to consider other aspects of travel behavior beyond travel frequency.

We review the literature to understand potential responses to disruptions and the factors that influence them. This helps us form expectations about how certain passenger groups might change their behavior, which we can later test. Additionally, studying these factors guides our selection of clustering indicators based on data availability.

2.1. Factors Influencing Passenger Responses to Disruptions

Passenger responses to disruptions can include changing modes, routes, departure times or stations, waiting for service to resume, or canceling journeys. These responses are influenced by various factors, including individual characteristics, journey specifics, network conditions, and the nature of the disruption. This section synthesizes findings from reviewed papers on how these factors influence passenger responses.

Travel Purpose

When the purpose of a journey is critical and urgent (e.g., work or school), passengers are more likely to change their routes (Adelé et al., 2019; Drabicki et al., 2021) or switch to alternative modes like cars and taxis (Nguyen-Phuoc et al., 2018b; Mo et al., 2022; Papangelis et al., 2016; Tian and Zheng, 2018; Nguyen-Phuoc et al., 2018a). They are less likely to cancel their journeys due to the mandatory nature and urgency of the trip (Rahimi et al., 2020).

Travel Duration

For longer trips, passengers are more likely to switch to cars or other modes (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018) due to increased uncertainty with public transport during disruptions (Rahimi et al., 2020; Nguyen-Phuoc et al., 2018b; Nguyen-Phuoc et al., 2018a).

Disruption Time (peak, off-peak)

Disruptions during peak hours increase the likelihood of passengers switching modes to ensure timely arrival at their destinations (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018). Passengers are more willing to spend money to reduce travel time and uncertainty during peak hour disruptions (Li, Yao, Yamamoto, Tang, & Liu, 2020).

Income and Employment

Higher-income individuals are more likely to switch modes, such as driving their own cars or using ride-sharing services, while lower-income groups, students, and the elderly tend to continue using public transport or shuttle buses (Li, Yao, Yamamoto, Huan, and Liu, 2020; Rahimi et al., 2020; Zhu et al., 2017; Saxena et al., 2019; Arslan Asim et al., 2021; Mo et al., 2022).

Travel Frequency

Frequent public transport users are more likely to continue using the system during disruptions due to their familiarity with the network, allowing them to find alternative routes more easily (Papangelis et al., 2016; Mo et al., 2022). These passengers exhibit lower sensitivity to disruption uncertainties (Li, Yao, Yamamoto, Tang, & Liu, 2020).

Availability of Alternative Routes

In areas with available alternative public transport routes, most demand can be accommodated by these alternatives (Rahimi et al., 2020). In suburban areas with limited public transport options, passengers are less likely to change routes and more likely to switch modes (Li, Yao, Yamamoto, Huan, and Liu, 2020; Mo et al., 2022; Adelé et al., 2019).

Disruption Duration

Short-term disruptions lead passengers to consider changing routes or modes or canceling their journey (Nguyen-Phuoc et al., 2018a). Longer disruptions increase the likelihood of switching to other modes (Li, Yao, Yamamoto, Huan, & Liu, 2020).

Disruption Frequency

In response to frequent disruptions, Drabicki et al. (2021) found that 77% of passengers made long-term adjustments, such as changing their routes or departure times to avoid disruptions. Similarly, Papangelis et al. (2016) reported that passengers switched modes, relocated, and changed jobs to cope with these disruptions.

Driver's License and Car Ownership

Passengers with a driver's license or access to a car are more likely to switch to driving during disruptions (Nguyen-Phuoc et al., 2018b; Adelé et al., 2019). Those without a license, such as students, are more likely to continue using public transport or cancel their trips (Nguyen-Phuoc et al., 2018a).

3. Methodology

To study long-term changes in passenger behavior due to disruptions, we first need to define aspects of this behavior, such as travel frequency and regularity, using a set of indicators. This allows us to measure and compare passenger behavior before and after a disruption and assess its impact. We have specific expectations about how certain passenger profiles might change their behavior; therefore, we use segmentation as it helps us identify and track these profiles over time. We implement a mixture latent Markov model, a probabilistic clustering algorithm, to identify travel pattern profiles in each period. By tracking how passengers transition between these patterns, we can observe changes in behavior, the extent of travel behavior inertia and other internal tendencies.

The analysis begins by selecting a suitable disruption and identifying the affected passengers. To ensure that behavior changes are directly attributable to the disruption, we control for external factors and seasonal trends by comparing affected passengers' behavior changes to those of a reference group.

3.1. Disruption Identification

To identify a suitable disruption several criteria are considered: (1) The disruption must occur on weekdays (excluding public holidays) and before 9 PM to primarily capture commuter behavior. (2) The disruption should last long enough to potentially have a lasting impact, such as several days/weeks at a station or multiple repeated shorter disruptions. (3) The Covid-19 pandemic period is excluded due to the significant decline in public transport usage. (4) There must be around one month prior to the disruption for the pre-disruption period, and a minimum of five months following the disruption for the post-disruption analysis. (5) There should be no other significant disruptions in the affected area during the pre- and post-disruption periods to isolate the disruption's impact.

Once a suitable disruption is identified, the pre-disruption period is determined to be approximately one month. The post-disruption period is then divided into intervals of roughly one month each, with public holidays excluded from these periods.

Automatic Vehicle Location (AVL) data is used to identify disruptions by analyzing the intervals between consecutive train departures. If these intervals are larger than usual, it indicates a potential delay at that station. To assess this, we compute two measures based on the departure time of trains:

- **Headway per line per direction:** The time between the departure of two consecutive trains at a station traveling in the same direction on the same line.
- **Headway per direction:** The time between the departure of two consecutive trains at a station traveling in the same direction, regardless of their line.

Whenever large values are observed for both of these two measures at the same time, we can conclude that a significant delay has occurred.

3.2. Identification of the Affected Passengers

Once a suitable disruption is identified, the next step is to filter out the passengers impacted by it. Frequent travelers who often start their journeys from the disrupted station are more likely to change their behavior due to their reliance on that station. Therefore, these frequent travelers are the focus of our analysis.

For a single event disruption, passengers with at least four journeys from the disrupted station during the pre-disruption period are considered affected passengers. For multiple disruptions, a two-step process is used: (1) Identify passengers who initiated at least four journeys from the affected stations around the time of the disruptions during the pre-disruption period. A specific time window is used to determine if a journey aligns with the disruption. This window is the interval ("headway") between the disrupted train's departure time and the preceding train's departure, plus an additional thirty-minute margin to account for variations, as illustrated in Figure A.1. A journey is considered frequent if its start time falls within the range [delayed train departure time - headway - 30 min, delayed train departure time + 30 min]. This "headway" parameter is the same as the *headway per line per direction* discussed in Section 3.1.

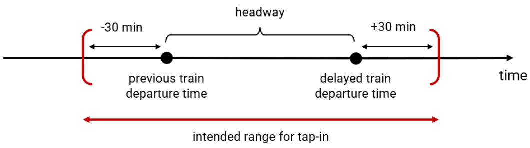

Figure A.1: The range for $tap\ in_i$ to consider i a frequent journey made around the disruption time.

(2) Among these frequent travelers, select those who experienced the disruptions at least twice. A journey is considered affected by a disruption if the tap-in time falls within [delayed train departure time - headway, delayed train departure time - scheduled headway]. This range is visualized in Figure A.2. This criteria ensures that people who tap-in just before the departure of the delayed train and do not fully experience the long delay are excluded. The scheduled headway is deducted from the departure time for this purpose because it is the usual interval between two trains on that line.

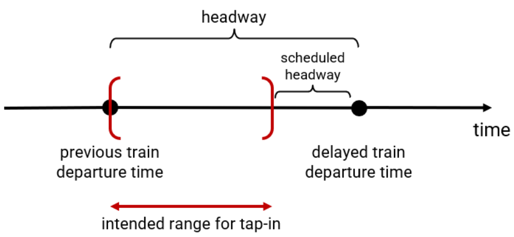

Figure A.2: The range for $tap\ in_i$ to consider i a journey affected by a disruption (delay).

3.3. Identification of the Reference Passengers

The reference group includes passengers who use a station similar to the disrupted one in terms of redundancy, regional characteristics (residential or workplace areas), and proximity to the city center. This group must remain unaffected by any disruptions to accurately reflect natural passenger behavior trends. Alternatively, reference passengers can be selected from the disrupted station if they were not impacted by the disruption. Passengers with at least four journeys during the pre-disruption period are chosen for the reference group.

3.4. Passenger Segmentation

In the literature, various clustering methods have been employed to study passenger travel behavior. This paper uses a variant of latent class analysis (LCA) known as the mixture latent Markov model (MLMM), which is suitable for clustering longitudinal data. This model allows individuals to transition between different latent *states* (travel patterns) over time (Vermunt & Magidson, 2013). Membership in these travel patterns is determined by several observed indicators capturing the different aspects of travel behavior. Model relationships are estimated via multinomial logit models and maximum likelihood estimation is used to determine model parameters (Magidson & Vermunt, 2002).

Individuals are assigned to latent *classes* based on their transition similarities between the states (Vermunt & Magidson, 2013). Inspired by the approach of Kroesen and van Cranenburgh (2016), we refer to each class as a "mobility style", reflecting the underlying attitudes that influence changes in travel patterns over time.

An advantage of MLMMs that motivated our choice is their ability to incorporate additional variables, often sociodemographic factors, known as covariates (Magidson & Vermunt, 2004). Although these variables are not part of the observed indicators, they can impact transition probabilities, travel patterns, and mobility styles. This capability is crucial for our study because whether a passenger is affected or reference may influence their likelihood of belonging to a specific mobility style and their transition probabilities. Therefore, "type" (i.e., whether a passenger is affected or a reference) is included in the model as a covariate, and its significance is subsequently tested.

A notable variant of MLMMs is the "mover-stayer" model, which distinguishes between individuals who remain in the same state (stayers) and those who change states (movers). This model is particularly useful for highlighting behavior stability and change within the population (Vermunt & Magidson, 2013). Additionally, separating the stayers helps better capture the heterogeneity among the movers (Kroesen & van Cranenburgh, 2016).

Model Assumptions

First, it is assumed that passengers' travel behavior is captured by a set of time-varying latent states, i.e., travel patterns, which at each period represent travel behavior with the smallest possible number of logit parameters. Six indicators are used to capture these latent travel patterns which are discussed in Section 3.4.1.

Secondly, we assume that a time-constant latent variable, referred to as mobility style, explains both the initial travel pattern membership and the probabilities of transitioning between different travel patterns over time. Mobility styles account for the heterogeneity in these transition probabilities.

Third, data points are clustered based on their similarity regarding the indicators, a process called the measurement model (Haustein & Kroesen, 2022) which is indicated in Figure A.3. These indicators are assumed to be mutually independent conditional on the latent state variable, known as the local independence assumption (Magidson & Vermunt, 2002). Furthermore, the part of the model which probabilistically assigns data points to the latent classes (mobility styles) is called the structural model (Haustein & Kroesen, 2022). This part of the model allows for the inclusion of covariates to explain class membership.

Finally, Figure A.3 shows that for every period $t \in \{0, ..., T\}$, the Markov model includes a latent state variable (travel pattern), and a multinomial logit model estimating relationships between successive latent state variables. This creates a matrix of transition probabilities, indicating the likelihood of transitioning to a future state based on the current state, following the first-order Markov assumption (Vermunt & Magidson, 2013).

3.4.1. Selected Indicators

We expect several changes in passenger behavior following a disruption, which guide our selection of indicators. Specifically, we expect reduced travel frequency, fewer travel days, and a decreased number of journeys, necessitating the definition of corresponding indicators. Additionally, the impact of disruptions may vary based on them happening during peak/off-peak hours and weekdays/weekends, altering passenger activity during these times. To avoid frequently disrupted hours, passengers might also adjust their usual travel times, leading to more variability in their daily travel start time. To test these expectations and capture all aspects of passengers' temporal travel behavior, the following indicators are defined and calculated for each passenger per period:

- **Share of active days:** The percentage of days during the period on which a passenger made at least one journey.
- **Average number of journeys per weekday:** The average number of journeys on weekdays.
- **Share of weekend journeys:** The share of journeys during the whole period which were made on Saturday and Sunday.
- **Share of peak-hour journeys:** The share of weekday journeys during the peak hours (6AM to 9AM and 3PM to 7PM) in a period.
- **Share of similar first boarding time:** The share of days a passenger started his/her first journey of the day around a similar time. This indicator enables the identification of travel time regularity and the extent to which passengers travel at the same time everyday which is a suitable measure of time flexibility of passengers.

  To calculate this indicator, an approach introduced by Bhaskar, Chung, et al. (2014) involves using DBSCAN to identify the densest areas in the dataset as groups of days with approximately similar boarding times. The tap-in time of the first journey of a day is converted into minutes-from-midnight. DBSCAN requires two parameters, ϵ and $MinPts$. $MinPts$ represents the minimum number of points needed to form a cluster, with clusters below this threshold classified as noise. $MinPts$ is typically set to twice the dimension (i.e., two in this case). ϵ indicates the maximum distance between two points to be considered neighbors.

  To determine the appropriate value for ϵ Sander et al. (1998) suggest using the ϵ of the smallest cluster. With $MinPts$ known, the first step is to calculate the distance of each data point to its kth nearest neighbor (where k is $MinPts$). These distances are then sorted and plotted, with the elbow of the resulting graph chosen as the value for ϵ. Points with distances larger than this value are labeled as noise, while the rest are assigned to clusters. The largest cluster size divided by the total number of active days for a passenger is set as the value for this feature.
- **Share of similar last boarding time:** This indicator is defined as the share of days a

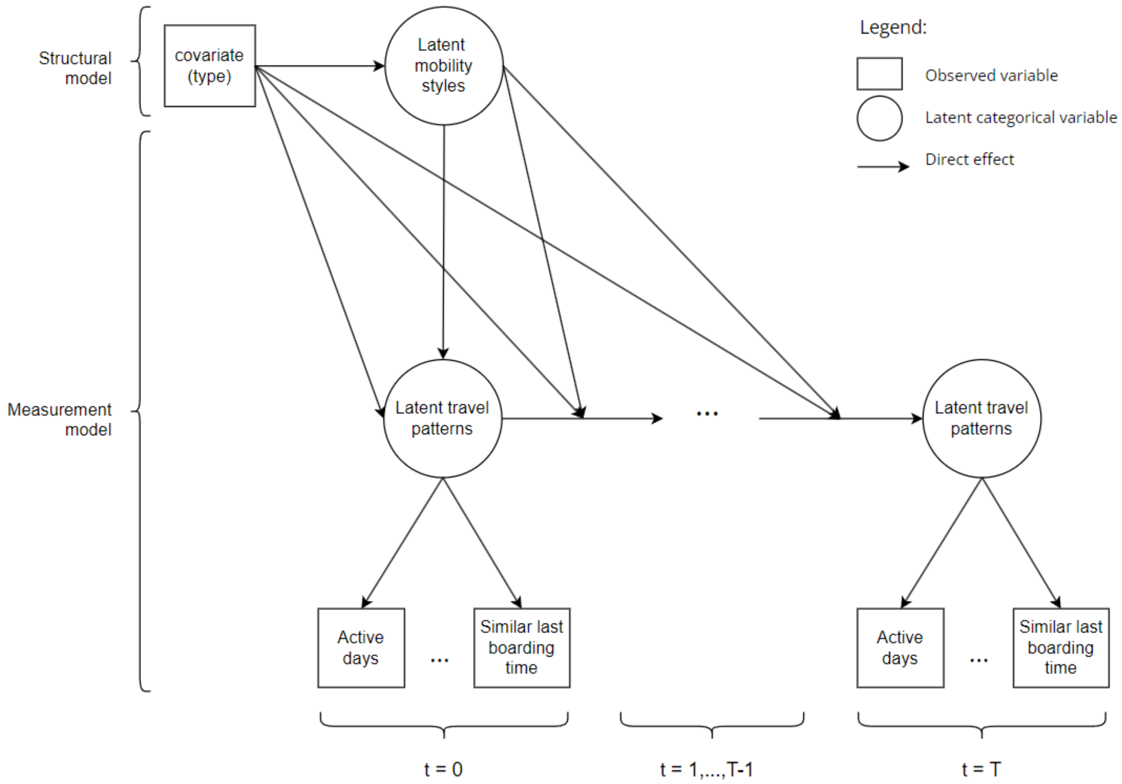

Figure A.3: Graphical representation of the latent class model.

passenger started his/her last journey of the day around a similar time. A similar approach to calculating the first boarding time is used to calculate this indicator.

4. Case Study

4.1. Case Study Description

The data used in this study is provided by the Washington Metropolitan Area Transit Authority (WMATA). The Washington D.C. metro network, known as Metrorail, consists of six color-coded lines as seen in Figure 4.4. The system design requires passengers to tap in their smart cards upon entering and exiting the network. The provided data includes Automated Fare Collection (AFC) data, Automatic Vehicle Location (AVL) data, and logs of planned and unplanned disruptions from August 2019 to December 2022.

The disruption identified in this study occurred in September, causing delays ranging from 20 to 37 minutes at five stations on the orange line towards Vienna. The locations of the affected stations are shown in Figure A.4, and the specific dates of disruptions, along with the total number of delay occurrences per station, are provided in Table A.1.

Next, the pre- and post-disruption periods were determined, as shown in Table A.2. These periods were selected to maintain a similar ratio of weekend days to the total number of days across all periods. Additionally, public holidays were excluded from the analysis for all periods.

4.2. Results

This section outlines the analysis procedure and presents the clustering results, divided into two sections. Section 4.2.1 discusses the selection of the optimal model, while Section 4.2.2 provides an analysis of travel patterns, transition probabilities, and mobility styles.

4.2.1. Model Definition

This section outlines the process of determining the optimal model, by explaining the selection of the optimal number of states and classes and the reasoning behind these choices. It also discusses the significance of the indicators and the covariate, and the resulting implications.

Optimal Model Selection

To determine the optimal number of travel patterns (states) and mobility styles (classes), a two-step approach based on the method by Kroesen and

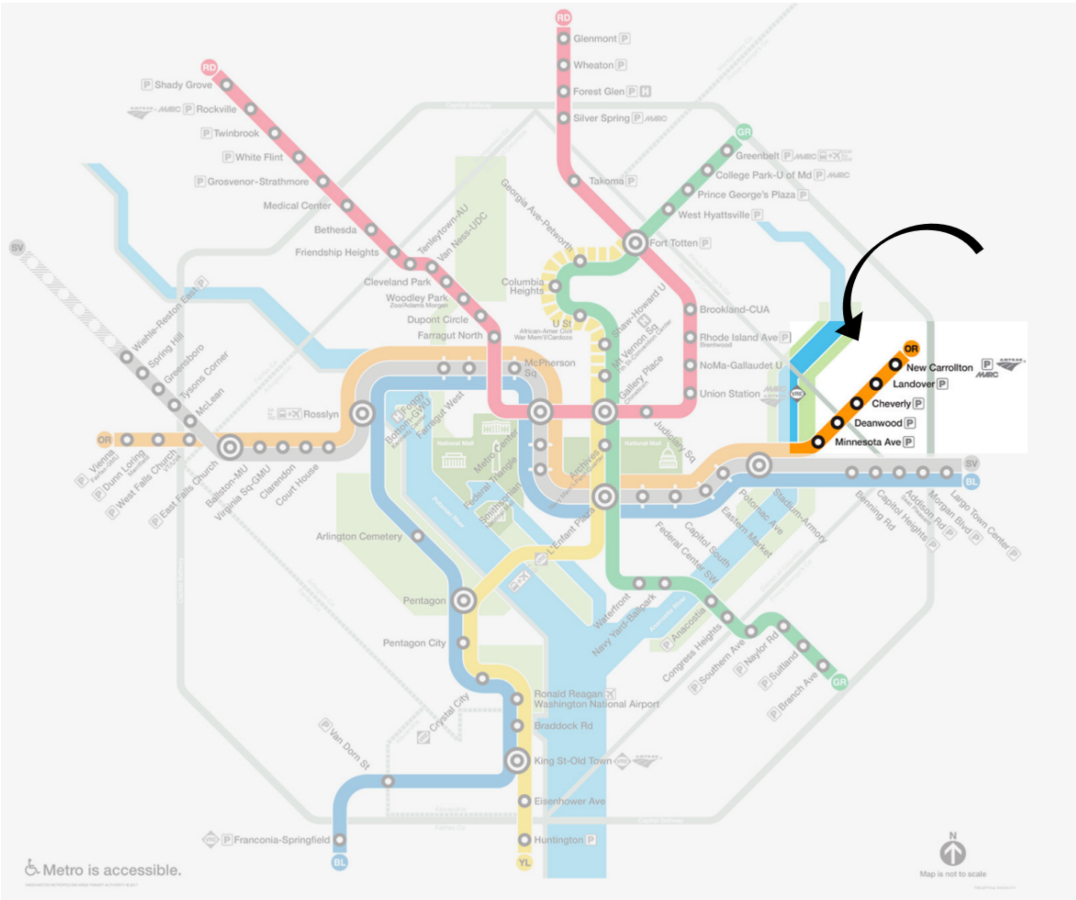

Figure A.4: Disrupted area identified for the analysis.

Table A.1: Disrupted stations and disruption dates.

Disrupted Station	Dates (in September 2019)	Number of Occurrences	Range
Minnesota Ave	9 - 11 - 19 - 20 - 23 - 24	11 times	20 to 37 min
Deanwood	11 - 19 - 20 - 23 - 24	8 times	20 to 33 min
Cheverly	11 - 19 - 20 - 23 - 24	9 times	20 to 32 min
Landover	19 - 20 - 23 - 24	7 times	20 to 32 min
New Carrollton	11 - 16 - 17 - 19 - 20 - 23 - 24	12 times	20 to 31 min

Table A.2: Pre- and post-disruption periods.

Period	Start	End	Duration
Pre-Disruption	2019-08-11	2019-09-07	27 days
Post-Disruption 1	2019-09-25	2019-10-26	31 days
Post-Disruption 2	2019-10-27	2019-11-26	30 days
Post-Disruption 3	2019-11-27	2019-12-21	24 days
Post-Disruption 4	2020-01-01	2020-02-01	30 days
Post-Disruption 5	2020-02-02	2020-02-29	29 days

van Cranenburgh (2016) is used. First, the optimal number of states is determined without considering covariates, which then is used for the selection of the number of classes.

Thus, first, LCA models are estimated, starting with one state and increasing up to ten states, with the optimal model chosen based on the Bayesian Information Criterion (BIC) (Nylund et al., 2007). The model with the lowest BIC value is selected as optimal. However, since the BIC value consistently decreases (as shown in Table A.3), and a ten-state model is impractical for a mixture latent Markov model (Kroesen & van Cranenburgh, 2016), an alternative approach is taken, prioritizing cluster interpretability and relevance. As a result, a five-cluster model is chosen, as beyond five clusters, the distinctions between clusters become less meaningful.

After determining the optimal number of states, a similar method is applied to identify the optimal number of classes. Mixture latent Markov models are estimated with five states and varying numbers of mobility styles, from one to four. Starting with models that include two classes, the ”stayer” class is incorporated. Although the model with ”1 stayer class and 3 mover classes” has the lowest BIC value (see Tabble A.4), the reduction in BIC compared to the ”1 stayer class and 2 mover classes” model is minimal, while the increase in the number of parameters is significant (21%). Consequently, the ”1 stayer class and 2 mover classes” model is selected for interpretation. The model is then re-estimated with the addition of the ”type” covariate.

Significance of the Indicators and the Covariate

The significance of the model's indicators was evaluated using the Wald test, revealing that all six indicators are significant, with p-values below 0.05. This supports their inclusion in the model, as they are important for distinguishing between travel patterns. Next, the covariate ”type” was added to the optimal model, and it was found to significantly influence initial state membership and transition probabilities (p-value < 0.05), but not class membership (p-value 0.27). This indicates that the covariate does not impact a passenger's likelihood of belonging to a specific mobility style, leading to the removal of the covariate's relationship with class membership from the model.

4.2.2. Analysis of the Results

This Section provides an explanation of travel pattern profiles, followed by the general transition probabilities between the patterns, which reveal the degree of travel behavior inertia along with the directions and extent of transitions from each pattern to the other over time. Additionally, the mobility styles and their characteristics are described, offering insight into the underlying tendencies among passengers. In this Section we assess the extent to which our expectations regarding the behavior change are met.

Travel Pattern Profiles

The travel pattern sizes (averaged over periods) and the mean value of the indicators for each of the five travel patterns are displayed in Table A.5. The information is displayed for the whole data, thus both the affected and reference. Using this information, we describe the characteristics of each travel pattern and provide a representative label for them.

- pattern 1 (size: 33%): **Frequent less flexible peak travellers**

  These passengers are classified as frequent travelers due to being active on 57% of days and averaging 1.53 journeys per weekday. They do not travel on weekends, as evident from the table and the graph, and have the highest proportion of journeys during peak hours (76%). Because their peak-hour activity exceeds 50%, we refer to them as peak travelers. Additionally, their boarding times show moderate flexibility compared to other groups, with 56% of days having a similar first boarding time and 49% having a similar last boarding time.

- pattern 2 (size: 27%): **Occasional less flexible peak travellers**

  These passengers are the second least active group, traveling on only 25% of days with an average of 0.62 journeys per weekday. Hence, they are referred to as occasional travelers. They do not travel on weekends and make 63% of their journeys during peak hours. Their boarding times indicate moderate flexibility.

- pattern 3 (size: 19%): **Frequent flexible peak travellers (with weekend)**

  These passengers are the most active group, traveling on 70% of days with an average of 1.55 journeys per weekday. They show an interest in weekend travel, accounting for 15% of their journeys, and 59% of their trips occur during peak hours. This group is the second most flexible in terms of boarding times.

- pattern 4 (size: 15%): **Occasional very flex-**

Table A.3: Model fit results.

No. of states (travel patterns)	LL	BIC(LL)	Param
1	-12973	26068	12
2	45460	-90667	25
3	57297	-114209	38
4	62975	-125434	51
5	67774	-134901	64
6	72203	-143628	77
7	75294	-149679	90
8	77585	-154129	103
9	79615	-158058	116
10	81721	-162139	129

LL = log-likelihood
BIC(LL) = Bayesian information criterion (based on log-likelihood)
Param = number of parameters
$BIC_{LL} = -2LL + \ln(samplesize)Param$

Table A.4: Model fit results.

No. of classes (mobility styles)	LL	BIC(LL)	Param
1 mover class	73801	-146904	84
2 mover classes	74488	-148069	109
1 stayer class 1 mover class	74200	-147660	89
3 mover classes	74655	-148195	134
1 stayer class 2 mover classes	74626	-148303	115
4 mover classes	74752	-148182	159
1 stayer class 3 mover classes	74760	-148363	139

LL = log-likelihood
BIC(LL) = Bayesian information criterion (based on log-likelihood)
Param = number of parameters
$BIC_{LL} = -2LL + \ln(samplesize)Param$

Table A.5: Profiles of the latent travel patterns (for the affected and reference passengers, averaged over all periods).

	1. Frequent less flexible peak travellers	2. Occasional less flexible peak travellers	3. Frequent flexible peak travellers (with weekend)	4. Occasional very flexible travellers (with weekend)	5. Sporadic travellers
Cluster size	33%	27%	19%	15%	6%
Indicators (mean)					
share of active days	0.57	0.25	0.70	0.33	0.04
journeys per weekday	1.53	0.62	1.55	0.55	0.07
share of weekend journeys	0.00	0.00	0.15	0.28	0.16
share of peak journeys	0.76	0.63	0.59	0.46	0.40
share of similar first boardings	0.56	0.47	0.44	0.29	1.00
share of similar last boardings	0.49	0.44	0.38	0.29	1.00

ible travellers (with weekend)

These passengers are active on only 33% of the days with 0.55 journeys per weekday and are thus called occasional travelers. They perform 28% of their journeys in the weekend. They perform 46% of their journeys during peak hours and are the second least active pattern during peak hours. They are the most flexible group with similar first and last boarding times of 29%.

- pattern 5 (size: 6%): **Sporadic travellers**

  These passengers are active on just 4% of days, with an average of 0.07 journeys per weekday. Thus, they are referred to as sporadic travelers. They make 16% of their journeys on weekends and 40% during peak hours. Additionally, they tend to start their first and last journeys at similar times each day.

Travel Pattern Transition Probabilities

Table A.6 shows the overall transition probabilities of the affected and reference passengers which shows how likely individuals are to move from one travel pattern to another between consecutive periods. The values which are substantially large (greater than 0.15) have been highlighted in bold. Below we provide a discussion about our observations based on this table, the expected outcomes, and the actual observed outcome. We interpret these observations to derive conclusions about the behavior changes of the affected passengers and their difference with the reference.

Travel Behavior Inertia:

This table shows that passengers in the three largest travel patterns (patterns 1, 2, and 3) tend to remain in the same pattern, with a likelihood of at least 66% among the affected group and 62% among the reference group. These three patterns contribute to almost 80% of the data (based on Table A.5). This large likelihood of remaining in the same travel pattern from each period to the next is an indication of travel behavior inertia among the majority of the passengers.

The highest travel behavior inertia is observed among passengers in the largest pattern, pattern 1 (Frequent less flexible peak travelers). These passengers have an 85% probability of maintaining their behavior from one period to the next in the affected group and a 77% probability in the reference. This is significant given that pattern 1 is the largest pattern in both groups.

Another insight is that among the affected group, the frequent travelers (patterns 1 and 3) are more likely to stay in the same pattern (85% and 66%, respectively) compared to occasional travelers of pattern 2 and 4 (67% and 36%, respectively). This indicates a strong travel behavior inertia among passengers with high travel frequency.

Travel Behavior Inertia: Affected vs. Reference:

We expected to observe lower travel behavior inertia (likelihood of staying in the same travel pattern) among the affected passengers compared to the reference group. This expectation was based on the idea that affected passengers, having experienced disruptions, would be more likely to modify their travel behavior. This expectation is only partially met, as the diagonal values among the reference group are larger compared to the affected passengers for most patterns (patterns 2, 4, and 5 which are the occasional and sporadic travelers). However, this is not the case for patterns 1 and 3, where the affected passengers exhibit higher values on the diagonal.

Travel Frequency:

We initially expected that disruptions might lead passengers to travel less frequently with public transport, possibly opting to work from home or use other modes of transportation like driving a car. Therefore, we anticipated high transition probabilities from frequent to occasional travel patterns. However, the majority of frequent affected travelers (patterns 1 and 3) remain in the same pattern, with probabilities of 85% and 66%, respectively. Additionally, 23% of passengers in pattern 3 shift to pattern 1, which is also a frequent travel pattern.

Peak-hour Activity:

Among the affected passengers, those who did not often travel during peak hours (pattern 4) are more likely to shift to patterns with higher activity during peak hours (patterns 1, 2, and 3) compared to the reference group. One possible reason for this is that the disruptions in our case study primarily occurred during off-peak hours. As a result, some passengers might adjust their travel timing to avoid potential disruptions similar to what they experienced during off-peak hours. Specifically, 62% of the affected passengers in pattern 4 (off-peak travelers) transition to the first three groups (peak travelers), whereas this value is only 30% for the reference group.

Travel Flexibility:

We observe that the patterns which are less flexi-

Table A.6: Transition matrices of the affected and reference passengers.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.849**	0.064	0.083	0.003	0.000
	2. Occasional less flexible peak travellers	**0.179**	**0.675**	0.018	0.095	0.035
	3. Frequent flexible peak travellers (weekend)	**0.235**	0.011	**0.656**	0.080	0.019
	4. Occasional very flexible travellers (weekend)	0.113	**0.367**	0.138	**0.356**	0.026
	5. Sporadic travellers	0.012	**0.609**	0.012	0.013	**0.355**
Reference	1. Frequent less flexible peak travellers	**0.772**	0.121	0.089	0.009	0.009
	2. Occasional less flexible peak travellers	0.091	**0.699**	0.026	0.094	0.091
	3. Frequent flexible peak travellers (weekend)	**0.174**	0.081	**0.618**	0.111	0.016
	4. Occasional very flexible travellers (weekend)	0.006	**0.234**	0.048	**0.612**	0.100
	5. Sporadic travellers	0.021	**0.292**	0.018	**0.165**	**0.504**

ble regarding their first and last boarding time (pattern 1 and 2) exhibit a higher travel behavior inertia compared to the patterns with higher flexibility (pattern 3 and 4). One possible reason might be because they have developed a habit of traveling at the same time every day (on the days that they use PT) and this habitual behavior might have extended to their overall travel pattern. Thus they tend to maintain the same travel behavior.

Overall Comparison of Affected and Reference Groups:

In conclusion, differences exist between the reference and affected passengers; however, the extent of this difference was expected to be more pronounced given the occurrence of the disruption. It is evident that this difference is less substantial for the first three patterns, which comprise 80% of the data. It is possible that if the disruptions were more significant, the differences would become more pronounced.

Analysis of the Mobility Styles

The passengers are divided into three mobility styles, based on their travel behavior changes over time. In this section the purpose is to use the transition probabilities of each of the travel patterns (Tables A.7, A.8, and A.9) for the affected and the reference, to provide a description for each mobility style. This helps us in identifying the different behavior traits that exist among the passengers (mobility styles).

Mobility style 1 (stayers):

The first mobility style is labeled as stayers, who stay in the same travel pattern across all periods and never change their behavior. Table A.7 shows that within this mobility style, the likelihood of staying in the same pattern is 1 as opposed to 0 for transitioning to other patterns.

Mobility style 2 (Commuter travellers):

This class is inclined towards travelling more frequently, less flexibly, less in the weekend, and more during peak. This inclination is more pronounced among the affected passengers than the reference:

- Higher frequency: Passengers tend to shift towards traveling more frequently. Based on Table A.8, frequent patterns (1 and 3), in both the affected and reference groups, are most likely to transition within the frequent patterns 1 and 3. This is due to the large transition probabilities of pattern 1 to 1 (80% among the affected and 67% among the reference), as well as pattern 3 to pattern 1 (82% among the affected and 63% among the reference) and pattern 3 to 3 (16% among the affected and 21% among the reference). Moreover, in the affected group, pattern 2 (the largest occasional pattern) is more inclined towards transitioning to the frequent pattern 1 rather than staying in the same pattern. Moreover, in the affected group, pattern 5 (sporadic travelers) shifts to pattern 2 (occasional travelers with higher frequency) with a 78% probability.
- Less flexibility and no weekend activity: Passengers tend to become less flexible and not travel on weekends. In both the affected and reference groups, the sum of the probabilities of shifting to less flexible patterns without weekend activity (1 and 2) is significantly higher than shifting to more flexible patterns with weekend activity (3 and 4).
- More peak-hour activity: Passengers are more likely to transition to patterns with higher peak-hour activity, particularly towards patterns 1, 2, and 3, with pattern 1 being the most active during peak hours.

Mobility style 3 (Leisure travellers):

This class is inclined towards travelling more flex-

Table A.7: Transition matrices of mobility style 1.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**1**	0	0	0	0
	2. Occasional less flexible peak travellers	0	**1**	0	0	0
	3. Frequent flexible peak travellers (weekend)	0	0	**1**	0	0
	4. Occasional very flexible travellers (weekend)	0	0	0	**1**	0
	5. Sporadic travellers	0	0	0	0	**1**
Reference	1. Frequent less flexible peak travellers	**1**	0	0	0	0
	2. Occasional less flexible peak travellers	0	**1**	0	0	0
	3. Frequent flexible peak travellers (weekend)	0	0	**1**	0	0
	4. Occasional very flexible travellers (weekend)	0	0	0	**1**	0
	5. Sporadic travellers	0	0	0	0	**1**

Table A.8: Transition matrices of mobility style 2.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.798**	0.124	0.077	0.001	0.000
	2. Occasional less flexible peak travellers	**0.490**	**0.426**	0.008	0.032	0.043
	3. Frequent flexible peak travellers (weekend)	**0.818**	0.011	**0.162**	0.004	0.006
	4. Occasional very flexible travellers (weekend)	**0.168**	**0.807**	0.001	0.023	0.002
	5. Sporadic travellers	0.052	**0.779**	0.002	0.002	**0.166**
Reference	1. Frequent less flexible peak travellers	**0.668**	0.203	0.108	0.006	0.015
	2. Occasional less flexible peak travellers	**0.161**	**0.649**	0.020	0.059	0.111
	3. Frequent flexible peak travellers (weekend)	**0.629**	0.134	**0.215**	0.011	0.011
	4. Occasional very flexible travellers (weekend)	0.014	**0.772**	0.001	0.070	**0.144**
	5. Sporadic travellers	0.035	**0.366**	0.006	0.057	**0.537**

ibly, less during peak, and more in the weekend while travelling with the same frequency. This inclination is more pronounced among the affected than the reference:

- Same frequency: Passengers tend to maintain their travel frequency. For example, frequent patterns (1 and 3) are more likely to transition within patterns 1 and 3. For pattern 1, the total transition probability to patterns 1 and 3 is 94% among the affected, and 84% among the reference. For pattern 3, the total transition probabilities to patterns 1 and 3 is 86% among the affected, and 65% among the reference. Similarly, the occasional patterns (2 and 4) majorly transition to occasional patterns (2 and 4). The sporadic travellers (pattern 5) either remain as sporadic travellers or transition to pattern 2 and 4 (occasional travellers) and thus increase their frequency slightly.
- More flexibility: Among the affected and the reference groups, passengers in pattern 1 (less flexible) are most likely to transition to pattern 3 (more flexible) rather than staying in the same pattern. The probability of the transition from pattern 1 to 3 is 58% among the affected and 61% among the reference group. Similarly, pattern 2 (less flexible) transitions to pattern 4 (more flexible) with a probability of 39% among the affected and 41% among the reference.
- Less peak-hour activity: Pattern 1 (most active pattern during peak) most likely transitions to pattern 3 which has less peak activity compared to pattern 1 and 2. Pattern 2 most likely transitions to pattern 4 which has even lower activity in peak. Finally, pattern 3 mostly remains as is.
- More weekend activity: We observe substantial transitions from patterns with no weekend activity (patterns 1 and 2) to patterns 3 and 4 which are active in the weekend.

Mobility Styles and Travel Patterns Sizes

The size of the mobility styles for the affected and reference passengers is presented in Table A.10. Comparing the probability of mobility style membership, i.e., total sizes, between the affected and reference passengers in Table A.10, we do not observe substantial differences. We expected the affected passengers to be less likely to be stayers compared to the reference. In other words, we anticipated a smaller class size for stayers among the affected passengers. Since these passengers have experienced the disruption, we expected them to be more likely to modify their be-

Table A.9: Transition matrices of mobility style 3.

		Travel pattern in period t				
	Travel pattern in period (t-1)	1	2	3	4	5
Affected	1. Frequent less flexible peak travellers	**0.359**	0.017	**0.582**	0.042	0.000
	2. Occasional less flexible peak travellers	0.041	**0.398**	0.065	**0.392**	0.104
	3. Frequent flexible peak travellers (weekend)	**0.164**	0.013	**0.690**	0.109	0.024
	4. Occasional very flexible travellers (weekend)	0.116	**0.373**	0.173	**0.335**	0.003
	5. Sporadic travellers	0.000	**0.727**	0.016	0.016	**0.241**
Reference	1. Frequent less flexible peak travellers	**0.225**	0.021	**0.611**	0.142	0.000
	2. Occasional less flexible peak travellers	0.008	**0.344**	0.088	**0.407**	**0.152**
	3. Frequent flexible peak travellers (weekend)	0.079	0.102	**0.573**	**0.217**	0.029
	4. Occasional very flexible travellers (weekend)	0.006	**0.209**	0.072	**0.599**	0.114
	5. Sporadic travellers	0.000	**0.191**	0.036	**0.337**	**0.436**

havior rather than continue traveling in the same manner as before. However, the percentage of stayers is 27% among the affected group and 24% among the reference, indicating that the likelihood of never changing behavior is almost similar for both groups.

Furthermore, the affected passengers are just as likely (46%) to belong to mobility style 2 (tendency towards traveling more frequently) as the reference group (46%). This indicates that the tendency towards traveling more frequently does not differ between the two group which contradicts our expectation that the affected passengers would be more likely to travel less frequently. We expected that this value would be significantly lower among the affected passengers compared to the reference. Regarding the third mobility style, which consists of passengers with a tendency towards off-peak travel, the class sizes are again quite similar with only 3% difference.

5. Discussion and Conclusion

In this study, we investigated the long-term effects of a public transport disruption in the Washington DC metro network on passenger travel behavior. Using a mixture latent Markov model, we identified five distinct travel patterns and monitored the transitions of passengers between these patterns over time.

The results reveal that the differences between the affected and reference passengers are not as significant as expected, particularly for the three largest travel patterns. This observation could indicate that the disruption did not have a substantial effect on changing passenger behavior. This might be because the disruptions had a low frequency and duration, leading to only temporary delays. This observation aligns with findings in the literature. For example, Drabicki et al. (2021) found that the higher the frequency of disruptions, the more likely it is to result in long-term behavioral changes among passengers. Moreover, in a study by Papangelis et al. (2016), passengers reported making permanent behavioral changes, such as mode changes, relocation, and job changes due to experiencing frequent long-term disruptions in the past.

Another reason for the minimal impact of the disruptions in this study could be that they primarily occurred during off-peak hours. Literature indicates that the timing of disruptions significantly influences passenger responses. During peak hours, passengers are more likely to change their travel mode, opting for taxis, buses, or bikes to avoid delays and ensure timely arrival (Li, Yao, Yamamoto, Huan, and Liu, 2020; Tian and Zheng, 2018). Li, Yao, Yamamoto, Tang, and Liu (2020) also found that passengers perceive travel time more negatively during peak hours and are willing to pay more to reduce it. Although these findings are based on immediate responses to disruptions, they may extend to long-term behavior if disruptions are frequent and repetitive, as noted by Drabicki et al. (2021) and Papangelis et al. (2016).

Furthermore, we observed a high level of travel behavior inertia among the affected passengers. Whereas we initially expected that passengers experiencing disruptions would be more likely to change their behavior, resulting in lower travel behavior inertia compared to unaffected passengers. In addition, we expected that the affected frequent travelers would shift to occasional travel patterns, and reduce their travel frequency due to the disruption. Instead, frequent travelers displayed high travel behavior inertia, continuing to travel as frequently as before. This observation aligns with the findings of Adelé et al. (2019) and Li, Yao, Yamamoto, Tang, and Liu (2020), who noted that frequent travelers are less likely to change their routes or modes when facing disruptions and are more likely to wait for the service to resume.

Several factors could explain this behavior. Fre-

Table A.10: Mobility style size of the affected and reference passengers.

	Total size	
Mobility style	affected	reference
1 (stayer)	0.267	0.238
2	0.462	0.456
3	0.272	0.307

quent travelers might be employees required to commute to the office daily without the option of working from home, or they might not own a car, limiting their alternatives. Consequently, they cannot reduce their use of public transportation even when experiencing disruptions. Additionally, their frequent use of public transportation might have become habitual, making them less sensitive to disruptions.

Regarding the sizes of mobility styles, we did not observe substantial differences between the affected and reference groups, which contradicts our expectations. Specifically, we anticipated a smaller size for mobility style 1 (stayers) among affected passengers, assuming they would modify their behavior due to disruptions and be less likely to maintain the same travel pattern. Lastly, the absence of a mobility style characterized by less frequent travel further indicates the minimal impact of the disruptions observed in our case study.

In conclusion, this study successfully develops a framework for studying the long-term impact of disruptions. However, the implementation of this framework in this study revealed that the case study disruption did not have a very substantial long-term impact.

5.1. Limitations and Future Research

One of the limitations of this study is a lack of detailed and accurate data on the start and end times of disruptions. This required inferring disruptions from train movements in the AVL dataset, which involved extensive data cleaning and special handling. Another limitation is the restricted effective period available for analyzing disruptions. The onset of COVID-19 in Washington, DC, in March 2020 limited our analysis to a seven-month period before the pandemic. Ideally, we would prioritize analyzing significant disruptions, such as station or line closures, over frequency reductions, as these are more likely to have a substantial long-term impact. However, such significant disruptions were not present in the available data for the case study period. Additionally, the unavailability of socio-demographic data was a limitation. Understanding the impact of factors such as income, age, car ownership, etc, on passengers' long-term travel behavior would be valuable.

Future research can extend the current methodology by incorporating socio-demographic information, such as age, income, education, etc., as covariates to evaluate their impact on transition probabilities, initial state membership, and class membership. Understanding the socio-demographic composition of each class can provide insights into why certain travel patterns emerge, such as preferences for peak versus off-peak travel. For instance, younger individuals might prefer off-peak hours due to flexible schedules, while older adults might travel during specific times for routine activities. These insights can inform targeted strategies to optimize transport services and address the specific needs of different user groups.

A second avenue for future research is analyzing changes in the spatial aspect of travel behavior in addition to the temporal aspect. Researchers can investigate whether the frequent boarding and/or alighting stations of passengers change after a significant disruption or if passengers use these stations less frequently. For example, when a regular boarding/alighting station is disrupted for an extended period, passengers might use other nearby, undisrupted stations instead. It would be interesting to determine whether passengers return to their frequent stations once the disruption is resolved, to what extent their usage levels recover, or if they continue using the new stations. Lastly, the framework can be applied to other case studies, particularly those involving more significant disruptions such as long-term station or line closures.

B

Appendix

Table B.1: Example of the dataset to be clustered.

CardID	Type	Period	Active Days	Journeys per Weekday	Peak-hour Journeys	Weekend Journeys	Similar First Boarding Time	Similar Last Boarding Time
ID001	disrupted	Pre-disruption	0.39	0.41	0.36	0.00	0.25	0.25
ID001	disrupted	post-disruption 1	0.53	1.52	0.00	0.80	0.12	0.71
ID001	disrupted	post-disruption 2	0.32	0.91	0.05	0.43	0.77	0.83
ID001	disrupted	post-disruption 3	0.72	1.94	0.75	0.43	0.22	0.56
...	...	...	...	...	...	...	...	...

Table B.2: Parameters of the mixture latent Markov model.

term			coef	s.e.	z-value	p-value	Wald(0)	df	p-value	Wald(=)	df	p-value
Class(1)	1		-0.30	0.04	-7.15	8.60E-13	86	2	1.80E-19			
Class(2)	1		0.35	0.04	8.55	1.20E-17						
Class(3)	1		-0.05	0.04	-1.16	0.24						
State[=0](1)	1	Class(1)	2.69	0.93	2.90	0.0037	494	12	4.00E-98	453	8	7.60E-93
State[=0](2)	1	Class(1)	1.45	0.93	1.56	0.12						
State[=0](3)	1	Class(1)	0.61	0.94	0.65	0.52						
State[=0](4)	1	Class(1)	-0.76	1.00	-0.76	0.45						
State[=0](5)	1	Class(1)	-4.00	3.64	-1.10	0.27						
State[=0](1)	1	Class(2)	3.27	0.96	3.41	0.00064						
State[=0](2)	1	Class(2)	1.88	0.96	1.95	0.051						
State[=0](3)	1	Class(2)	0.30	0.97	0.31	0.76						
State[=0](4)	1	Class(2)	-1.25	1.03	-1.21	0.23						
State[=0](5)	1	Class(2)	-4.21	3.76	-1.12	0.26						
State[=0](1)	1	Class(3)	0.94	0.88	1.07	0.29						
State[=0](2)	1	Class(3)	0.88	0.88	1.00	0.32						
State[=0](3)	1	Class(3)	1.83	0.88	2.08	0.038						
State[=0](4)	1	Class(3)	0.61	0.93	0.66	0.51						
State[=0](5)	1	Class(3)	-4.26	3.41	-1.25	0.21						
State[=0](1)	type		-0.25	0.93	-0.28	0.78	25	4	4.40E-05			
State[=0](2)	type		0.42	0.93	0.45	0.65						
State[=0](3)	type		1.17	0.94	1.25	0.21						
State[=0](4)	type		1.61	0.98	1.64	0.1						
State[=0](5)	type		-2.94	3.63	-0.81	0.42						
State(1)	1	Class(2) State[-1](1)	0.00	.	.	.	948	40	8.80E-173	803	20	3.90E-157
State(2)	1	Class(2) State[-1](1)	-1.86	0.27	-6.87	6.40E-12						
State(3)	1	Class(2) State[-1](1)	-2.34	0.29	-7.95	1.80E-15						
State(4)	1	Class(2) State[-1](1)	-6.39	1.23	-5.21	1.80E-07						
State(5)	1	Class(2) State[-1](1)	-9.07	7.42	-1.22	0.22						
State(1)	1	Class(2) State[-1](2)	0.14	0.55	0.26	0.8						
State(2)	1	Class(2) State[-1](2)	0.00	.	.	.						
State(3)	1	Class(2) State[-1](2)	-3.93	1.06	-3.70	0.00021						
State(4)	1	Class(2) State[-1](2)	-2.58	0.65	-3.98	6.80E-05						
State(5)	1	Class(2) State[-1](2)	-2.30	0.63	-3.64	0.00028						
State(1)	1	Class(2) State[-1](3)	1.62	0.37	4.38	1.20E-05						
State(2)	1	Class(2) State[-1](3)	-2.69	1.05	-2.56	0.011						
State(3)	1	Class(2) State[-1](3)	0.00	.	.	.						
State(4)	1	Class(2) State[-1](3)	-3.84	0.90	-4.26	2.10E-05						
State(5)	1	Class(2) State[-1](3)	-3.38	1.07	-3.14	0.0017						
State(1)	1	Class(2) State[-1](4)	1.99	1.69	1.17	0.24						
State(2)	1	Class(2) State[-1](4)	3.56	0.94	3.78	0.00015						
State(3)	1	Class(2) State[-1](4)	-3.35	7.23	-0.46	0.64						
State(4)	1	Class(2) State[-1](4)	0.00	.	.	.						
State(5)	1	Class(2) State[-1](4)	-2.52	7.47	-0.34	0.74						
State(1)	1	Class(2) State[-1](5)	-1.16	8.91	-0.13	0.9						
State(2)	1	Class(2) State[-1](5)	1.55	3.45	0.45	0.65						
State(3)	1	Class(2) State[-1](5)	-4.65	8.04	-0.58	0.56						
State(4)	1	Class(2) State[-1](5)	-4.69	7.98	-0.59	0.56						
State(5)	1	Class(2) State[-1](5)	0.00	.	.	.						
State(1)	1	Class(3) State[-1](1)	0.00	.	.	.						
State(2)	1	Class(3) State[-1](1)	-3.04	1.37	-2.22	0.027						
State(3)	1	Class(3) State[-1](1)	0.48	0.39	1.23	0.22						
State(4)	1	Class(3) State[-1](1)	-2.16	1.13	-1.90	0.057						
State(5)	1	Class(3) State[-1](1)	-11.88	10.48	-1.13	0.26						

State(1)	1	Class(3) State[-1](2)	-2.27	1.07	-2.12	0.034			
State(2)	1	Class(3) State[-1](2)	0.00	.	.	.			
State(3)	1	Class(3) State[-1](2)	-1.82	1.05	-1.73	0.083			
State(4)	1	Class(3) State[-1](2)	-0.02	0.65	-0.02	0.98			
State(5)	1	Class(3) State[-1](2)	-1.34	0.64	-2.09	0.037			
State(1)	1	Class(3) State[-1](3)	-1.44	0.33	-4.30	1.70E-05			
State(2)	1	Class(3) State[-1](3)	-3.95	1.01	-3.90	9.50E-05			
State(3)	1	Class(3) State[-1](3)	0.00	.	.	.			
State(4)	1	Class(3) State[-1](3)	-1.85	0.39	-4.80	1.60E-06			
State(5)	1	Class(3) State[-1](3)	-3.36	0.72	-4.65	3.30E-06			
State(1)	1	Class(3) State[-1](4)	-1.06	0.91	-1.16	0.25			
State(2)	1	Class(3) State[-1](4)	0.11	0.68	0.16	0.87			
State(3)	1	Class(3) State[-1](4)	-0.66	0.87	-0.76	0.45			
State(4)	1	Class(3) State[-1](4)	0.00	.	.	.			
State(5)	1	Class(3) State[-1](4)	-4.89	7.43	-0.66	0.51			
State(1)	1	Class(3) State[-1](5)	-6.61	11.47	-0.58	0.56			
State(2)	1	Class(3) State[-1](5)	1.10	3.37	0.33	0.74			
State(3)	1	Class(3) State[-1](5)	-2.71	7.95	-0.34	0.73			
State(4)	1	Class(3) State[-1](5)	-2.70	7.95	-0.34	0.73			
State(5)	1	Class(3) State[-1](5)	0.00	.	.	.			
State(1)	type	State[-1](1)	0.00	.	.	.	51	20	0.00018
State(2)	type	State[-1](1)	0.67	0.27	2.47	0.014			
State(3)	type	State[-1](1)	0.52	0.30	1.72	0.085			
State(4)	type	State[-1](1)	1.70	1.14	1.49	0.14			
State(5)	type	State[-1](1)	5.25	7.42	0.71	0.48			
State(1)	type	State[-1](2)	-1.53	0.54	-2.84	0.0046			
State(2)	type	State[-1](2)	0.00	.	.	.			
State(3)	type	State[-1](2)	0.46	1.05	0.44	0.66			
State(4)	type	State[-1](2)	0.18	0.63	0.29	0.77			
State(5)	type	State[-1](2)	0.53	0.63	0.84	0.4			
State(1)	type	State[-1](3)	-0.55	0.35	-1.55	0.12			
State(2)	type	State[-1](3)	2.22	1.02	2.18	0.029			
State(3)	type	State[-1](3)	0.00	.	.	.			
State(4)	type	State[-1](3)	0.88	0.40	2.22	0.026			
State(5)	type	State[-1](3)	0.36	0.75	0.49	0.63			
State(1)	type	State[-1](4)	-3.63	1.03	-3.52	0.00043			
State(2)	type	State[-1](4)	-1.16	0.68	-1.71	0.088			
State(3)	type	State[-1](4)	-1.46	0.88	-1.66	0.097			
State(4)	type	State[-1](4)	0.00	.	.	.			
State(5)	type	State[-1](4)	3.24	7.43	0.44	0.66			
State(1)	type	State[-1](5)	-1.59	8.91	-0.18	0.86			
State(2)	type	State[-1](5)	-1.93	3.42	-0.56	0.57			
State(3)	type	State[-1](5)	0.20	7.97	0.03	0.98			
State(4)	type	State[-1](5)	2.44	7.96	0.31	0.76			
State(5)	type	State[-1](5)	0.00	.	.	.			
share active days	1		0.38	0.00	236.27	1.1e-12124	55,822	1	1.1e-12124
share active days	State(1)		0.19	0.00	104.47	1.1e-2372	113,111	4	9.9e-24558
share active days	State(2)		-0.13	0.00	-54.55	1.2e-648			
share active days	State(3)		0.32	0.00	121.40	3.5e-3203			
share active days	State(4)		-0.05	0.00	-11.81	3.50E-32			
share active days	State(5)		-0.34	0.00	-212.92	3.5e-9847			
journeys per weekday	1		0.87	0.00	207.37	9.3e-9341	43,001	1	9.3e-9341
journeys per weekday	State(1)		0.67	0.01	129.49	1.1e-3643	76,431	4	6.3e-16593
journeys per weekday	State(2)		-0.25	0.01	-40.56	1.4e-359			
journeys per weekday	State(3)		0.69	0.01	83.73	6.1e-1525			

journeys per weekday	State(4)	-0.31	0.01	-42.51	7.4e-395			
journeys per weekday	State(5)	-0.80	0.00	-181.94	4.2e-7191			
share weekend journeys	1	0.12	0.00	40.07	5.1e-351	1,605	1	5.1e-351
share weekend journeys	State(1)	-0.12	0.00	-40.07	5.1e-351	12,534	4	1.1e-2718
share weekend journeys	State(2)	-0.12	0.00	-40.07	5.1e-351			
share weekend journeys	State(3)	0.04	0.00	11.01	3.50E-28			
share weekend journeys	State(4)	0.16	0.00	36.34	3.80E-289			
share weekend journeys	State(5)	0.04	0.01	3.86	0.00011			
share peak journeys	1	0.57	0.00	149.16	4.5e-4834	22,248	1	4.5e-4834
share peak journeys	State(1)	0.19	0.00	43.33	3.3e-410	2,619	4	2.3e-566
share peak journeys	State(2)	0.06	0.01	11.11	1.10E-28			
share peak journeys	State(3)	0.02	0.01	3.82	0.00013			
share peak journeys	State(4)	-0.11	0.01	-16.77	4.30E-63			
share peak journeys	State(5)	-0.16	0.01	-12.31	8.00E-35			
similar first boardings	1	0.55	0.00	403.65	2.5e-35383	162,930	1	2.5e-35383
similar first boardings	State(1)	0.01	0.00	2.36	0.019	115,625	4	1.0e-25103
similar first boardings	State(2)	-0.08	0.00	-26.82	1.80E-158			
similar first boardings	State(3)	-0.11	0.00	-31.88	4.70E-223			
similar first boardings	State(4)	-0.26	0.00	-87.93	7.5e-1682			
similar first boardings	State(5)	0.45	0.00	328.45	7.5e-23429			
similar last boardings	1	0.52	0.00	400.10	8.2e-34765	160,082	1	8.2e-34765
similar last boardings	State(1)	-0.03	0.00	-11.78	4.80E-32	137,729	4	3.0e-29903
similar last boardings	State(2)	-0.08	0.00	-27.22	3.30E-163			
similar last boardings	State(3)	-0.14	0.00	-43.01	3.7e-404			
similar last boardings	State(4)	-0.23	0.00	-74.19	5.5e-1198			
similar last boardings	State(5)	0.48	0.00	368.65	1.2e-29513			