Data-driven customer segmentation: Assessing disparities in COVID impact on public transit user groups and recovery

by
Rachel Li-Jiang Luo
A.B. in Economics, Harvard University, 2013

Submitted to the Department of Urban Studies and Planning
in partial fulfillment of the requirements for the degree of
Master in City Planning, Master of Science in Transportation
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2021

Author . . . .
Department of Urban Studies and Planning
May 14, 2021

Certified by . . . .
Jinhua Zhao
Edward H. and Joyce Linde Associate Professor
Thesis Supervisor

Certified by . . . .
Jim Aloisi
Lecturer of Transportation Policy and Planning
Thesis Supervisor

Certified by . . . .
Joanna Moody
MITEI Mobility Systems Center Research Program Manager
Thesis Supervisor

Accepted by . . . .
Ceasar McDowell
Professor of the Practice
Chair, DUSP MCP Committee

Data-driven customer segmentation: Assessing disparities in COVID impact on public transit user groups and recovery

by

Rachel Li-Jiang Luo

Submitted to the Department of Urban Studies and Planning
on May 14, 2021, in partial fulfillment of the
requirements for the degree of
Master in City Planning, Master of Science in Transportation

Abstract

COVID-19 triggered an unprecedented global lockdown and severely dampened public transit ridership, which was down 62% year-on-year across the U.S. through Q4 2020 [1]. Beyond these stark headline figures, more granular views of *whose* transit ridership patterns changed and *how* are needed to aid cash-strapped transit agencies in understanding both the operational and equity impacts of COVID-19 and assessing possible recovery strategies [2].

This thesis examines these questions in the Metro Boston region by applying k-means clustering to smart card data from the Massachusetts Bay Transportation Authority (MBTA). We empirically determine customer segments based on passenger-level pre-pandemic transit ridership patterns during January 13 - February 16, 2020, using data from 22.6 million trips by 1.5 million passengers. We then trace how COVID-19 produced differential churn rates and travel behaviour modifications among these distinct passenger groups. We find that COVID-19 induced churn among rail commuter segments key for supporting MBTA fare revenues, while bus riders and those who frequently rode rail off-peak—groups that covered the majority of reduced-fare and vulnerable passengers—were most likely to continue using the system.

Our findings suggest that in the near term, the MBTA can support a ridership and revenue rebound by working closely with large employers involved in the MBTA "Perq" corporate pass program to plan for reopening. This can also position the MBTA to better gauge the need to redesign or reprice Perq to offer greater flexibility for workers who may be adopting remote work longer term and therefore commuting less frequently to the office. Further, our analysis reveals consistency in ridership patterns among bus passengers even during crisis times. In the medium term as the MBTA considers network redesigns to meet post-pandemic travel needs, existing plans for bus upgrades do not necessarily need heavy modification because COVID-19 did not completely redefine these passengers' transit usage patterns. This gives a base level of certainty for the MBTA's planning process, as it seeks to track and shape the uncertainty that COVID-19 has brought to demand on the rail side of its network. Finally, by supplementing our quantitative analysis with an overview of COVID-19 responses by other major U.S. transit agencies, we suggest that the MBTA can better weather future emergencies like COVID-19 by making longer-term efforts to shift its operating revenue mix away from volatile fare revenues towards more stable and resilient revenue sources such as sales and property taxes, and complementing this with sustainable financial management.

The framework offered in this thesis for dissecting passenger ridership behavior and tracking passenger churn and cluster-switching can be applied to other transit agencies to detail either background ridership behavioral changes in normal years or rapid step-changes during a mobility crisis. Understanding passengers' ridership demand at the cluster level can inform both immediate actions that transit agencies can take to enable recovery, as well as support network redesign and long-term resilience.

Thesis Supervisor: Jinhua Zhao
Title: Edward H. and Joyce Linde Associate Professor

Thesis Supervisor: Jim Aloisi
Title: Lecturer of Transportation Policy and Planning

Thesis Supervisor: Joanna Moody
Title: MITEI Mobility Systems Center Research Program Manager

Acknowledgments

This thesis and the research behind it could not have been possible without the support of Professor Jinhua Zhao, Professor Jim Aloisi, and Dr. Joanna Moody from the JTL-Transit Lab and MIT Energy Initiative Mobility Systems Center (MITEI MSC). Your careful guidance, your commitment to equitable and sustainable transportation, and the positive environment you built at the lab helped me to thrive during my time at MIT.

All members of the JTL-Transit Lab and MITEI MSC—past and present—were also integral to the work that went into this thesis, by providing feedback, new angles of research, and encouragement. I am especially indebted to Ruben Morgan, Qingyi Wang, Mary Rose Fissinger, Xiaotong Guo, Patrick Meredith-Karam, Hui Kong for the inspiration and support I found in their complementary work on COVID-19's mobility impact and on decision trees. I am also grateful to Randy Field from MITEI MSC for providing sharp insights that helped to keep my presentations of this work focused. Yunhan Zheng and Annie Hudson, thank you for introducing me to JTL during my first year at MIT, and for encouraging me to apply for the MST joint degree working for the lab. Further, I am thankful to Yunhan and Shenhao Wang for encouraging and accompanying me in the exploration of data science methods applied to transportation, and to Xin Yang for her constant support on our green mobility policy work. Lastly, I can't move on from JTL-Transit Lab without thanking the members of the Student Advisory Committee (SAT).

This work would also not have been possible without the MBTA's partnership with the Transit Lab. In particular, I am deeply indebted to Ian Thistle at the MBTA for helping me to understand the ODX and AFC SQL databases and guiding me through how best to use this data for clustering analysis.

At DUSP more broadly, I give my thanks to Professor Siqi Zheng, Professor David Hsu, and Professor Lawrence Susskind for all their advice during my years here, which have changed my worldview profoundly. At MIT writ large and beyond—thanks to all my friends at the MIT Graduate Christian Fellowship for providing me a north star and a community when I was tired and wondering where I was headed. At home, I am grateful to my husband Winston for helping me laugh when I was stressed, and to the nearby Longfellow House and Fresh Pond with their plentiful animal wildlife and flowers. To my parents and my parents-in-law, who let me move back in with them for COVID-19 summers/winters—it is a joy to have you in my life. To my father-in-law in particular—thank you for your encouragement of my transportation studies, pulling from your own PhD years working on the topic. To my undergraduate friends who stuck with me through thick and thin—thank you for always being there, and being excited when I try to tell you about my research with JTL-Transit Lab.

Contents

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

1.1 Background and motivation

The World Health Organization declared COVID-19 a global pandemic on March 11, 2020, accelerating government attempts to limit mobility and enforce social distancing as preventive measures [3]. The resulting business closures and the mass adoption of remote work for office workers led to a constriction of urban mobility and plummeting transit ridership around the globe. The United States alone logged 19.6 million COVID-19 cases by December 31, 2020 and saw public transit ridership fall 62% year-on-year by the fourth quarter [4, 1]. This nosedive was already an improvement from April 2020, when stay-at-home orders were rolling out across the U.S. and pulling transit usage down by as much as 81% year-on-year.

Alongside the downward spiral in ridership came falling fare revenues, coupled with a simultaneous shrinkage in other common but economy-dependent transit agency revenue sources such as sales taxes [5]. Uncertainty about the trajectory of the virus, the length of lockdowns, and the possibility of government stimulus created further difficulties for transit agencies attempting to adjust service to both meet the remaining demand and steward their finances. Newspaper headlines announced "apocalyptic" conditions facing the survival of public transit and asked "Will Mass Transit Recover From the Pandemic?" [6, 7]. Meanwhile, new studies seemed to indicate a positive association between transit usage and the risks of COVID-19 infection in places like

Manhattan [8, 9].

Over one year later at the start of May 2021, 45% of the U.S. population has received at least one vaccine dose and mobility appears to be returning across the country [4]. Yet transit still faces an uncertain path to recovery. Study after study indicate a seismic shift in employees' acceptance of remote work. A McKinsey analysis of 2,000 work tasks across 800 jobs and nine countries suggested that over 20% of the workforce can effectively work from home three to five days a week [10]. A Federal Reserve Bank of Atlanta survey found as of May 2020, firms anticipate up to 30% of their workforce to work from home at least once a week compared to only 10% pre-pandemic [11]. With the office commute potentially becoming less of a weekday fixture, it is becoming increasingly questionable whether public transit ridership patterns and volumes will be able to recover to pre-pandemic levels—or if recovery will instead require re-shaping transit services. The future remains up for debate and will be defined by negotiations between stakeholders including employers, their employees, government agencies, and regulators. Yet, transit agencies are beginning to plan for a robust recovery even in the face of uncertainty.

The roll-out of smart cards and automated fare collection (AFC) data in recent years provides transit operators with a tool to capture the evolution of transit ridership behavior in real time with trip-level granularity, for the entire passenger population [12]. During a sudden crisis like COVID-19, it also provides an established data collection platform that requires little additional adjustment to provide information during the crisis. As such, it can be a valuable complement to traditional surveys, which will still remain valuable as they capture passenger profile details such as socio-economics, access to alternative modes, and trip purposes. AFC data can be used to identify distinctive travel behavior patterns and categorize passengers into these rider "segments." Further, as the literature review in Chapter 2 will discuss, the application of clustering methods to identifying these passenger segments can aid transit agencies in operational and network planning.

This thesis applies a clustering approach to AFC data in order to give a granular view of key transit usage patterns for a case study transit agency, the Massachusetts

Bay Transportation Authority (MBTA) in the Metro Boston area. The continuation of AFC data through most of the COVID-19 pandemic is then used for tracking and dissecting the differential rates of churn and behavioral shifts among each major passenger segment. As transit agencies look ahead towards coordinating a ridership, service, and financial recovery, understanding the divergent ways in which their varied passenger segments were affected by COVID-19 can help them chart more robust and equitable paths towards recovery.

1.2 Research aims

This thesis aims to provide a framework for partitioning transit riders into behavioral segments or clusters based on smart card data regarding each rider's interactions with the transit system. Attention is paid to conducting clustering through passenger features interpretable for policy-making, and in applying other machine learning techniques to understand the ways in which passenger features are being used by the clustering algorithm to identify major underlying rider groups and sort individuals into them. The framework seeks to produce clusters that capture distinctive demand patterns being exerted upon the system, which can inform transit agency network design and operations.

To exploit the value of AFC data for behavior tracking in emergency situations such as the pandemic, this thesis also uses clustering to analyze how the composition of distinctive demand patterns shifted once the pandemic hit Metro Boston. Additionally, by tracking passengers who already existed in the baseline pre-pandemic dataset through the pandemic period, this thesis is able to trace which behavioral clusters were associated with higher churn rates, and how members of one cluster modified their behavior during the pandemic and thereby entered other clusters. AFC data has also been continuously available in Metro Boston for the years leading up to COVID-19 and was collected in a methodologically consistent way, which allows for comparison of the pandemic-period ridership changes to "background" changes observed in non-pandemic years, thereby helping to isolate COVID-19 ridership impacts from underlying trends.

Smart card data also carries meta-data on the fare product type and whether it is associated with reduced fares or other special fares like corporate passes. By examining the meta-data profile of each cluster, this thesis demonstrates how such additional smart card information can be used to create a more socio-economically rich profile of the types of riders that tend to fall into each behavioral cluster—and how these may relate to pandemic transit usage.

Finally, this thesis considers the COVID-19 response of the MBTA alongside that of 12 other domestic transit agencies, and combines this comparison with the quantitative clustering and churn analysis to provide short-, medium-, and long-term recommendations for recovery planning.

In this policy section and also throughout the thesis, we discuss transit equity from several angles including servicing reduced-fare riders and the differential impact of COVID-19 across behavioral clusters. However, our analysis does not entail an overarching synthesis of equity in transit planning and pandemic recovery because it focuses mostly on describing ridership patterns pre-COVID, tracking behavioral shifts during the pandemic, and comparing transit agencies' pandemic responses. There is room for future work to take up this important line of inquiry.

1.3 Thesis organization

This thesis begins with a literature and background review in Chapter 2, which covers market segmentation in the business settings where the concept originated, as well as how it has been applied to the public transit sector in particular. The chapter then focuses on the usage of clustering methods for transit passenger segmentation before giving an overview of recent studies published on the impact of COVID-19 on transit ridership. It finishes by summarizing the pandemic's effects in the Metro Boston area and familiarizing readers with the MBTA's bus and subway/trolley network, commonly called the "T."[1]

Chapter 3 gives an overview of the AFC derivative data source, called ODX, that we use for this thesis. It then reviews the k-means method and associated work necessary for running a clustering analysis, including data pre-processing, the creation of policy-relevant features as candidates for clustering, automated criteria for selecting among those features, and hyperparameter selection. Further, it lays out the additional steps we took to ensure temporal robustness or stability of the model, since we are applying it over time to track cluster evolution. The chapter then explains the optimal classification tree (OCT) methodology used to interpret how k-means created the final clusters. Finally, the chapter closes with a discussion of the data used for socioeconomic profiling of the clusters, our churn analysis methodology, and policy analysis.

Chapters 4 and 5 present our baseline and pandemic-era results, respectively. This includes interpretation of baseline pre-COVID clusters using OCT, and profiling with both smart card meta-data and separate, more aggregate MBTA survey data to enrich our understanding of the *who* within each cluster. Chapter 4 further dissects the operational and equity relevance of each cluster to MBTA planning. Chapter 5 follows by laying out the divergent ways in which each cluster responded to the pandemic.

The thesis closes with Chapter 6, which assesses the MBTA's COVID-19 response in light of our findings and actions taken by U.S. and international transit agencies facing similar crises. It provides immediate and longer-term priorities for consideration by the agency as it aims for a robust recovery in the face of ample uncertainty.

Last but certainly not the least, we note that in this thesis, the pronoun "we" is used in acknowledgement and thanks for the advice and guidance received from the committee, other members of the JTL-Transit Lab, and our data providers at the MBTA.

[1]This thesis does not analyze the MBTA's commuter rail ridership, for which trip records are stored in a separate database.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Literature Review

Understanding how riders use public transit is key to transit agencies' ability to deliver quality service, operate efficiently, and generate revenue. The advent of AFC and computational methods for handling these types of high volume, real-time data opened up new ways for transit operators and researchers to unpack the behavioral patterns of public transit riders, simulate transit networks, and improve origin-destination estimation algorithms [13, 14]. The granularity and timeliness of such data provides support for transit agencies planning across a range of scales, from daily operations to longer-term network design.

Unlike traditional survey-based studies, AFC is also updated live by the existing fare system, offering constantly refreshed views on the evolution of transit usage. Further, while traditional surveys may or may not be comparable over time depending on sampling methodology, AFC data captures the entire population of riders and is more likely to be consistent over time, giving better ability for monitoring ridership and revenue [12]. This automatic longitudinal tracking and comparability is especially valuable for understanding behavioral shifts when exogenous shocks hit a transit system, such as what urban areas around the world experienced with the COVID-19 pandemic in 2020.

For these reasons, an increasing number of studies have explored AFC to track the mobility impact of COVID-19 and its potential equity implications. AFC is often combined with other data sources such as the U.S. census for these purposes, since

AFC by itself does not capture passenger socio-demographics [12]. By extrapolation from census tracts adjacent to transit route stations and stops, such case studies in cities across the U.S. have confirmed popular intuition that greater ridership loss occurred in areas with higher-income households, and high shares of workers whose jobs transitioned well into a work-from-home environment. Meanwhile, census tracts dominated by lower-income and essential workers tended to continue using transit.

This thesis expands on the existing literature by applying k-means clustering to AFC-derived data for the MBTA, in order to identify separate groups of passengers who followed certain sets of transit usage patterns and therefore exerted diverging demand pressures on the system pre-pandemic. Following this market segmentation analysis, we trace each passenger cluster through COVID-19, to assess major demand shifts and how they reflect mobility adaptations riders were making (or unable to make) in response to the pandemic. By combining temporal features of the AFC-derived data with modal, spatial, and user type information collected by the same source, we profile each of these key behavioral clusters in terms of their modal affinities and their connections to revenue generation. This also allows us to give preliminary indications of socioeconomic differences in COVID-19's ridership impact.

The rest of this chapter introduces market segmentation in terms of its broader industry usage and its applications to public transit. Then, it will describe the clustering approach to transit market segmentation before reviewing existing work on COVID-19's transit impact across a variety of cities. Finally, it will give an overview of the emerging body of work on COVID-19's effects on mobility, and our case study area of Metro Boston.

2.1 Passenger segmentation for public transit

2.1.1 Market segmentation

The concept of segmentation is not unique to passengers in the transit domain. The field of market segmentation was first established by Wendell Smith in 1956, who noted that neoclassical economic theories of perfect competition and pure monopoly did not fully capture the realities of markets where competition is in fact imperfect [15]. Smith noted that while perfect competition requires homogeneous markets, demand for products is in fact heterogeneous and each market can be disaggregated into smaller, more homogeneous sub-markets. The original goal of such segmentation was to partition the demand side of a market to a scale where marketing and product differentiation efforts can be adjusted to target the particular needs and tastes of each sub-group of relatively more homogeneous customers. This is in contrast to mass marketing, where large populations of potential customers are all provided the same standardized marketing or product mix.

Market segmentation's success as a business tool has led to its expanded application in contemporary times, where firms "identify, profile, target, and reach segments using their own customer transaction data-bases" for product and marketing customization, making this a particularly useful method for customer retention [16]. Depending on the consumption habits and profile of a customer segment, businesses can choose certain competitive strategies such as lowering costs or creating a more premium, differentiated product. Greater access to granular customer transactions data and more sophisticated computational and statistical techniques have now extended the customization spectrum from mass marketing, through customer segmentation, to one-to-one tools targeting the individual.

However, companies operating in heterogeneous markets have not flocked directly to one-to-one tactics, recognizing that segmentation into major customer groups is often a relevant scale at which to operate in order to capture economies of scale in production, logistics, or marketing [16]. Further, segmentation provides a scale that has proven useful for strategy development by managers, even in cases where implementation of the strategy involves one-to-one tools. For market segmentation to be effective for these cases, researchers have argued that segments need to 1) have *measurable* characteristics along which they can be distinguished, 2) be *substantial* enough in size and profitability to serve, 3) capture customer segments actually *accessible* or *reachable* for the business, 4) be clearly *differentiable*, i.e. members are similar to each other but clearly different from those of other segments, and 5) *actionable*, giving rise to a basis for operational or product strategy [17].

Market segmentation faced by a particular business is generally empirically determined. A variety of techniques have evolved for this task including mixture models, hierarchical Bayes', clustering, and neural networks, and they have been applied to a wide range of industries from tourism and beverages to transportation [18, 19, 20, 16]. Clustering methods are popular among such studies. Chiu *et al* (2009) and Huang, Zeng, and Ong (2007) applied clustering to the education and beverages respectively. Kuo *et al* (2002) used neural networks to generate self-organizing feature maps upon which to run clustering analysis [20]. In transportation, the most famous colloquial discussion of market segmentation breaks down transit users into "captive" riders who depend upon the system due to a lack of alternatives, and "choice" riders who have viable alternatives yet choose to travel by transit—but further decompositions of transit riders have been undertaken by a number of studies, several of which use AFC as the "customer transaction record" found in industry [19].

2.1.2 Application of market segmentation to public transit

Since AFC data became more widely available, researchers have applied clustering to explore patterns in transit rider behavior and inform service enhancements by transit authorities. These studies also seek to use *measurable* features of their passengers' interactions with the transit system to identify *differentiable* customer sub-groups that allow for *actionable* outcomes for transit planning.

However, market segmentation for transit provision also differs from business applications. In contemporary urban areas, transit is subsidized and regulated in recognition of its role as a public good that provides access to economic and social opportunities while reducing the congestion, energy use, and air pollution associated with travel [21, 22]. Further, it is subsidized out of recognition that transit and its positive externalities have economies of scale and density [23]. As such, unlike in other market segmentation cases, we believe the goal of passenger market segmentation here is not to better drive profit as in the private sector, but rather to improve operational and strategic decisions that allow the transit system to meet its goals of providing access and sustainability [14]. Thus, more nuance is needed in discussions of a "substantial" customer segment—segments that are smaller but important for a transit system's equity goals can, from a public policy perspective, merit attention from planners despite its smaller size.

We note though that segmenting passengers based on smart card transaction records—that is, rider *behavior* on the system—may help us to discover ways in which groups typically considered "small" actually fit together into larger homogeneous segments. Transit systems often provide reduced-fare products for groups such as students, seniors, the blind, and those with disabilities, which are expected to be relatively smaller shares of the total ridership compared to full-fare adult riders. This manual segmentation of rider groups does not, however, prevent some reduced fare riders from having similar transit demand as other groups. This is one motivation behind the structure of this particular thesis, which focuses on customer segmentation based on behavior due to the operational relevance of this approach, then uses fare user types to "profile" the behavioral clusters produced. In this respect, the thesis follows Briand *et al* (2017).

Transit planners working with equity goals in mind may also approach the concept of "reachable" in market segmentation differently than private businesses, because of public transit's role in reaching those currently without transit access. That said, assessments of how to extend service to under-served passenger segments (or those currently not served) requires data separate from the smart card transaction records that are the core of this study, since by definition there is little transactions data on the potential transit usage patterns of those who currently can rarely ride, if at all. This aspect of achieving public transit's equity and sustainability goals may find more methodological commonality with the business sector's work in reaching new customer pools potentially with new products, rather than market segmentation of existing customers based on transactions data.

Finally, while transit market segmentation should strive to produce actionable results just as when segmentation is done for private enterprises, the types of actions under consideration can be significantly different. Some response strategies can be shared, such as offering reduced fares to certain segments or providing differentiated service (e.g. higher frequency service during peak hours for those who generally use transit only for commutes) [16]. However, the degree to which pricing and service provision can be changed are subject to regulatory and policy decisions that have shaped system design. For example, current discussions about moving towards a free transit system in our case study area of Metro Boston targets equity goals but would, if passed, remove pricing as an actionable lever that can be customized to each passenger segment [24]. The MBTA is also obligated under the Americans with Disabilities Act to provide service to those with disabilities, which it does through its paratransit service (RIDE) as well as TAP cards for those with disabilities who are still able to physically use the existing transit system [25]. If a particular passenger segment happens to be small but dominated by these riders, the MBTA is obligated by law to meet the needs of these riders.

Overall, there are several major transit agency goals that can be supported by passenger segmentation analysis [26]:

- **Evaluation and strategy for service improvements that enhance retention:** Knowing the major categories of passenger travel needs can inform efforts to cost-effectively enhance service quality and drive rider retention.

- **Travel demand management:** Passenger segmentation can help transit agencies target their travel demand management and incentive schemes efforts towards users with specific types of operationally relevant behavior patterns.

- **Evaluating impact of a major supply- or demand-side "shock":** This includes understanding the impact of a crisis like COVID-19 on system ridership, operational needs, and revenue; it further includes assessing the potential impact of a network expansion, route change, or introduction of a competitor service on ridership and user retention.

- **Marketing and product differentiation:** As in typical business applications of market segmentation, this type of analysis can be used by transit agencies to target different types of fare products (e.g., peak period pricing, commuter passes) and to optimize placement of revenue-generating advertisements on transit vehicles and facilities.

- **Informing survey sampling procedures:** By distinguishing how passengers are heterogeneous over space, modes, and time, passenger segmentation can help inform more balanced sampling techniques for transit ridership surveys.

Because the focus of this particular study is on understanding the variegated impacts of the COVID-19 crisis on the MBTA's heterogeneous ridership and informing a robust service recovery, we focus on the evaluation of a demand-side shock and how customer segmentation can inform recovery and service during a crisis.

2.1.3 Clustering approaches to transit market segmentation

Clustering as a transit passenger market segmentation methodology is generally built off AFC data, unlike traditional segmentation analysis using surveys [27, 14]. Such unsupervised methods are popular in these settings where there is no available ground-truth "label" to learn for passenger or trip classification. Though AFC enables individual-level analysis, this is often too specific and noisy, while clustering based on temporal and spatial habits offers a more actionable scale for network planning [14].

A variety of clustering methods have been used for passenger segmentation, using the temporal aspects of smart card data, the spatial, or both. Hierarchical clustering and k-means are common, as well as DBSCAN and Gaussian mixtures. Ghaemi *et al* (2017) applied hierarchical clustering to temporal data of passengers entering the public transit network, to classify users into behavioral groups based on the timing of their system entries [28]. Ma *et al* (2013) applied DBSCAN to trip chains to detect key historical travel patterns at the individual passenger level; it also used k-means on several extracted features—number of travel days, number of similar first boarding times, number of similar route sequences and number of similar stop ID sequences—with rough-set theory to cluster riders according to their usage regularity [29]. Kieu *et al* (2015) used k-means on trip chains to discover clusters of infrequent versus frequent riders, then put frequent riders' trip chains through DBSCAN to give a more granular clustering based on their boarding and alighting times and locations [27]. In Morency *et al* (2006) k-means was again used to identify clusters with similar boarding times based on the temporal profile of each passenger's boardings [30]. Meanwhile, Lathia *et al* (2013) used hierarchical clustering on the daily temporal profiles of each passenger's trip set to categorize rider behaviors and target information services [31].

Clustering based on temporal features has also been combined with other smart-card level meta-data to offer a fuller "profile" of riders in each cluster that can help transit agencies aiming to improve service for certain transit-dependent groups. Briand *et al* (2017) applied Gaussian mixture models to group passengers based on their temporal habits of transit ridership, identifying 10 rider clusters representing for example regular afternoon ridership, riders with two-peak temporal profiles, and riders with weekend morning activity [14]. They then profiled the clusters using card types, finding for example that students and seniors tended to fall into the cluster with a three-peak weekday commuting pattern—implications that can help transit agencies plan for the transit usage needs of the vulnerable and transit-dependent.

Meanwhile, El Mahrsi *et al* (2017) clustered smart card data from the complementary perspective of station operations and passenger flows, using a generative model-based approach to promote interpretability. On the station side, the paper clustered on number of transactions at the station by hour and day, which allowed the model to distinguish between stations with mostly balanced usage despite some rush hour peaks, and other stations with unbalanced usage. On the passenger side, temporal travel profiles of relatively frequent users were fed into a mixture of unigrams model. The model identified four clusters involving various sorts of diffuse temporal patterns for transit usage, dominated by young subscribers, those traveling on free passes, and the elderly. Two clusters exhibited typical commuting behavior and were mostly composed of regular paying subscribers, while four more were commuters with a secondary peak on Wednesday midday reflecting the school schedule of young riders.

Among MIT theses, Basu (2018) applied principal component analysis (PCA) to temporal and spatial features from AFC data from the Hong Kong Mass Transit Railway (MTR), then k-means clustered riders based on the main features that emerged from PCA [26]. His analysis was conducted for the MTR's information targeting campaign, in which the agency sought to better deliver system information to passengers for whom the updates were relevant. The interpretability of features was therefore of less importance in his analysis, hence his choice to use PCA rather than hand-crafted features with policy content.

Separately, Fissinger (2020) applied k-means clustering to AFC data from the Chicago Transit Authority, first to understand customer segments based on interpretable, hand-crafted features in normal conditions, then extending the analysis to COVID-19 passenger pattern changes [32]. The clustering portions of this thesis is most closely tied to Fissinger (2020) and Briand *et al* (2017), focusing on clustering using temporal AFC data but using spatial and smart card meta-data features for subsequent clustering profiling. It introduces the MBTA's operations in Metro Boston as a new case study of this type of passenger segmentation analysis.

2.1.4 Evolution in passenger segments over time

Fewer studies tracked the evolution of customer behavioral segments over time. Basu (2018) calculate correspondence scores for how stable cluster membership was between two years, finding that for the MRT clusters were relatively stable over the medium term, limiting the need to constantly update and maintain clusters [26]. Fissinger (2020) assessed the degree to which members of each cluster churned during COVID-19, discovering that free riders cluster and frequent off-peak bus users were most likely to continue riding during the pandemic, while frequent rail riders mostly churned [32]. Linking to pass types, these findings confirm that disadvantaged riders were the most likely to continue using the system during the pandemic.

Briand (2017) assessed temporal stability of cluster partitions over five years, to inform transit network operators regarding the level of maintenance necessary from year to year [14]. It found that riders were most likely to stay in the same cluster over time, or move to other clusters most similar to it as measured by Kullback-Leibler divergence. Morency *et al* (2006) used k-means clustering on temporal patterns of boarding to identify regular travel behaviors among transit users in Gatineau, Quebec. Subsequently, the paper looked at individual cards whose meta-data indicated different sociodemographic groups—a regular pass and an elderly pass—comparing the time series of their boardings over 277 consecutive days using k-means to identify groups of days exhibiting homogeneous behavior within each socio-demographic group [30].

This study's structure most resembles Briand's work in 1) clustering baseline ridership patterns from temporal AFC data, 2) profiling riders of each cluster using sources including smart card meta-data, and 3) conducting longitudinal analysis to track churn and cluster evolution. Further, our use of correspondence scores to assess cluster stability between seasons and years draws most closely from Basu (2018), and our focus on the churn impact of COVID-19 parallels Fissinger (2020).

This thesis contributes to the literature by tying these pieces together to give a fuller picture of the impact of a severe public health and economic crisis on usage patterns across socio-edemographic groups. The remaining sections of this literature review gives background on the COVID-19 crisis and its mobility effects, before zooming in on the case of the Metro Boston area serviced by the MBTA. This will provide context for our clustering and churn analysis as well as the policy recommendations they inform.

2.2 COVID-19's impact on transit ridership

COVID-19 prompted urban lockdowns and social distancing protocols across the globe, severely dampening public transit ridership. In the U.S., public transit ridership is still down 55% through the end of April 2021 compared to mid-February 2020, directly before the pandemic's presence became clear in the country [1].

Several studies have applied spatial analysis to smart card and U.S. census data to track transit usage during COVID-19 and link it to socio-demographics in order to understand its social implications [33, 34]. Liu, Miller, and Scheff (2020) applied logistic and regression analysis to data from the Transit app to describe the rate of transit ridership declines across the U.S. and their socioeconomic linkages [2]. Hu and Chen (2021) applied Bayesian structural time series models and regressions to Chicago smart card data to establish the impact of the pandemic on ridership and tie decline severity to census tract socio-economics [35]. Brough, Freedman, and Phillips (2020) leveraged automated passenger counts (APC), census, and survey data to study disparities in travel intensity declines in King County, Washington [36]. These studies found greater ridership loss in areas with higher-income, white, well-educated households; non-physical jobs translatable to remote work; and commercial real estate.

Public health and econometric studies using census data have also tied lower-income and essential worker status to continued transit use [9, 37]. McLaren (2020) specifically found a positive correlation between the share of African Americans and First Nations peoples in a county, and the area's total COVID-19 deaths. He further claimed that public transit use explained a statistically significant share of the variation in total pandemic deaths [9]. Sy *et al* (2020) conducted a spatial analysis at the zip code level that suggested essential worker status, which is tied to lower socio-economic conditions, was the risk factor with the strongest COVID-19 association. The study further suggested that essential workers' higher need to travel by subway was associated with higher COVID-19 case rates per 100,000 people, though this association became weaker after controlling for a zip code's median income [37].

Our study builds on this emerging literature by contributing new passenger-level analysis that is only made possible through access to transit agency AFC data. We offer 1) analysis based on each passenger's actual historical transit usage patterns rather than socio-demographic inference from adjacent census tracts; 2) a different approach to understanding customer behaviour using clustering and interpretation of clustering results through decision trees, and 3) a case study using MBTA smart card data and a system-wide passenger survey with subway station-level and bus route-level characteristics of actual riders.

2.3 Metro Boston during COVID-19

In response to COVID-19, Massachusetts entered a state of emergency on March 10, 2020, initiating a transit ridership freefall [38]. By August 2020 when fare collection data fully resumed, transit usage was on average still stagnating at 28% of February volumes. Figure 2-1 plots the state's daily case load as recorded by the Center for Disease Control against Google location history data on how visits to transit stations, workplaces, and retail/recreation changed compared to a January 3 - February 6, 2020 baseline. The data shows that even when caseloads were low over summer and the so-called "Phase II" stage of re-opening had commenced with the re-introduction of limited retail, dining, and office work, transit ridership was below pre-pandemic levels. Transit's usage fluctuated even as workplace visits appeared to have stabilized at a below-normal level, pointing to the role that choice work commuters and other trip purposes such as shopping and leisure affect transit demand during this time.

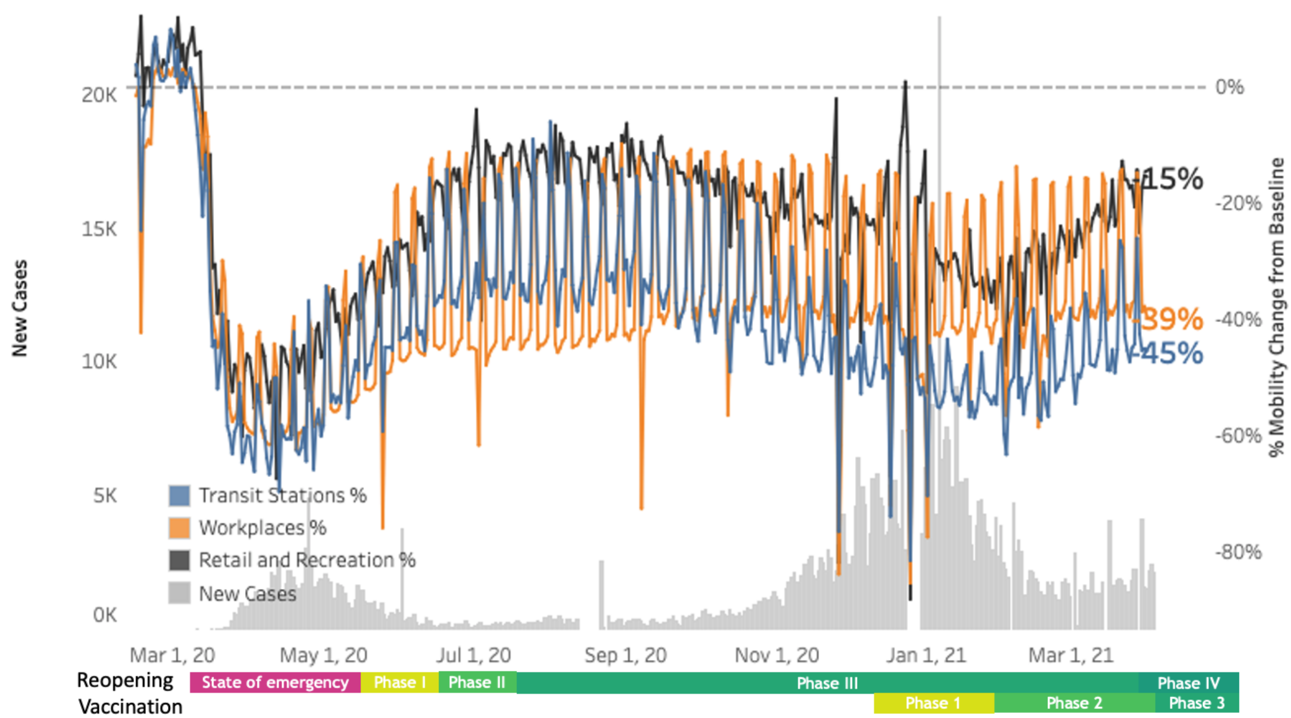


Figure 2-1: Massachusetts daily caseload and mobility patterns

Figure 2-2 focuses in on the the Boston region and compares how key travel modes used in the area—transit, walking, driving—evolved over the course of the pandemic

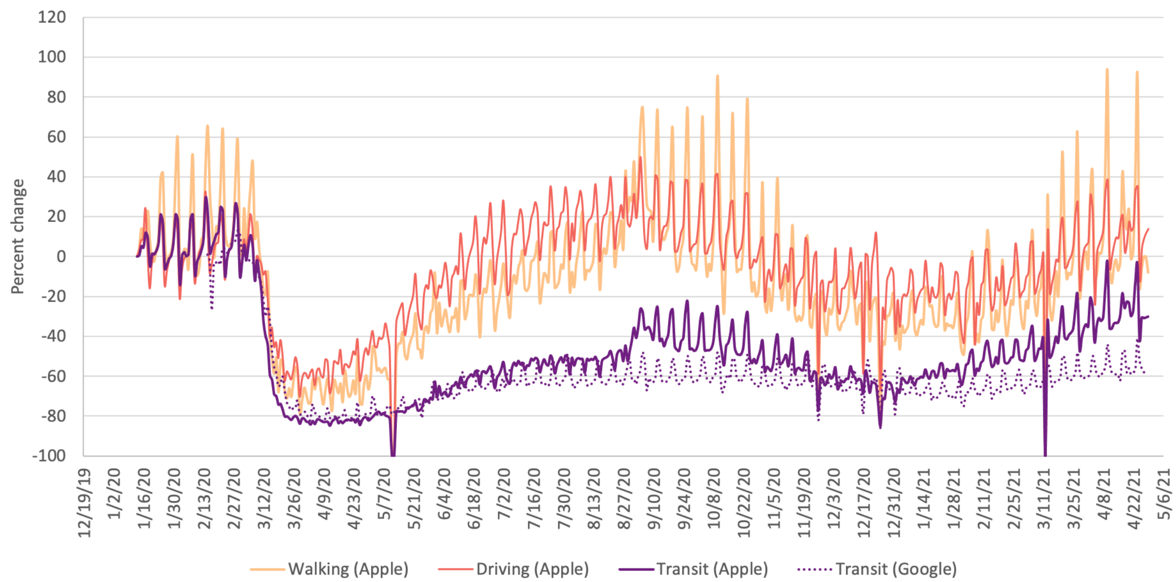


Figure 2-2: Boston area mobility patterns (walking, transit, driving)

compared to a pre-COVID baseline.[1] The data reveals the sharp fall of as much as -81% in transit use by April 2020 and a slow recovery that has not yet returned to pre-pandemic levels. This is in contrast to more private modes of transportation like walking and driving. Walking saw a sharp rebound to as high as 91% *above* pre-pandemic conditions by October 2020; driving rebounded as high as 50% above baseline levels by the start of September 2020. Transit usage did not begin a steady recovery until after vaccinations began in December 2020 (with a 1.5 month or so gap to account for the time it takes for a person to receive a second dose then become fully immune). News sources and the MBTA's own analysis tracked these trends as well as the impact it had on the agency's budget and ability to provide services [39, 40].

The aggregate figures do not, however, tell us how the transit usage decline and recovery occurred in different parts of the MBTA's extensive system, called the "T." The T is composed of a subway and bus rapid transit component (orange in Figure 2-3 below) that reflects a hub-and-spoke design, with dense service in downtown Boston and thinner service in outlying parts that are instead filled in by bus (purple in the same figure). Bus covers 8,047 stops and train service 166 stations including locations

[1] January 13 for Apple-derived time series, January 3-February 6, 2020 for transit data from Google

above and below ground; however, trips on rail outstrip bus trips on both weekdays and weekends. Separately, the MBTA operates a commuter rail service which extends outside of Metro Boston and is not considered in this study.

Outside of divergence in geographic reach between bus and rail, there are other clear areas of ridership divergence in parts of the system that could suggest differences in behavior during COVID-19. The MBTA's 2015-2017 System-wide Passenger Survey found, for instance, that 17% of bus riders surveyed during that period were reduced-fare pass users, compared to 10% on rail. At the route level, there is even greater variation, with route 24 and 33 in South Boston showing 44% reduced fare pass users. Bus riders were also, overall, 9 ppt less likely to have a household vehicle compared to rail users (30% versus 39%), with the share carless running as high as 54% on route 28 which goes through the Roxbury neighborhood. Thus, a more granular analysis of passenger behavior and its evolution during COVID-19 is needed to break down aggregate ridership trends to a level granular enough to provide insight for operational decisions like service cuts and strategies for robust recovery in the face of severe revenue shortages. Further, an approach incorporating temporal information to passenger segmentation can be a useful operational support to existing MBTA surveys, which only capture the spatial dimension.

2.4 Implications for this study

COVID-19 caused unprecedented upheaval to urban transit systems around the globe, crushing ridership and revenues and leaving uncertainty as to the timeframe and the shape of the subsequent recovery. This study builds on past works that tracked passenger clusters over time to study the crisis at a scale granular enough to inform operational and strategic planning for recovery [26, 14, 32, 30]. The pandemic was a period that revealed the transit-dependent rider groups and their variegated transit usage patterns, while simultaneously revealing riders who had the ability to opt out of the system. Passenger segmentation in the baseline and tracking throughout the pandemic and recovery therefore provide one way of deepening the MBTA's under-

Figure 2-3: MBTA bus system, rail + bus rapid transit system

standing of the specific behavioral habits and mobility needs of its riders, as well as the multi-pronged approach necessary to support each segment's recovery.

Based on existing techniques, we leverage temporal and modal AFC features to segment passengers into what we will call rider "clusters," to align with the name of the unsupervised learning technique we are leveraging to create the segmentation. Because the goal is to create clusters interpretable for transit planners and policy makers, we hand-craft potential features to be used in clustering rather than rely oon less interpretable feature creation processes like PCA. Based on findings from Briand (2017), we also recognized that there is background churn and cluster switching among passengers even in non-crisis times. This will inform our use of temporal stability/robustness checks prior to our COVID-19 analysis, so as not to confuse natural churn and cluster-switching with the impact of COVID-19 on these same behaviors. Finally, following existing literature, we leverage smart card meta-data and spatial data to profile each cluster, which gives the additional depth necessary for us to use our quantitative results to provide qualitative policy recommendations that can meet the needs of MBTA's diverse ridership while supporting the restoration of revenues.

Chapter 3

Data and Methodology

3.1 Overview

The growing availability of automated fare collection (AFC) data has enabled the application of machine learning methods to the detection of transit usage patterns and their evolution over time. Additionally, since AFC collects data at the level of the individual passenger (or, more strictly, the individual transit card or ticket), it offers greater granularity for the analysis of temporal and geospatial ridership patterns than was previously available with aggregated systems data and survey results.

As described in the literature review, existing analyses of COVID-19's impact on public transit ridership have begun to leverage these data sets, usually in combination with U.S. census or American Community Survey data. Regression and spatial analysis methods are applied to infer the sociodemographic composition of AFC trip records based upon the demographic profiles of adjacent census tracts and relate this to ridership volume changes during COVID-19 [36, 33, 2, 34].

In contrast with these existing papers, this thesis analyzes rider behaviour during COVID-19 at the passenger unit rather than at the geographic unit. AFC first allows us to assess pre-pandemic transit rider behavior by using individual passengers' pre-pandemic baseline transit usage to categorize them into major behavioral clusters—each of which creates a distinct set of demands on the transit system. Then, AFC allows us to track how these clusters, and the associated demands they each create for the system, evolve during COVID-19. Thus, applying behavioral clustering on AFC data can support transit agencies in making evidence-based planning decisions for the future travel demands of their customers, based on historical rider behavior collected by the agencies' ticketing system.

Clustering also provides analysis at a level of granularity relevant to transit agency operations. Briand *et al.* (2017) pointed out that working only at the individual rider level introduces too much noise into transit analysis; by contrast, cluster-level aggregation can be useful for providing a unit of analysis with enough aggregated volumes to describe extensive transit ridership activity on a network. That is, clustering strikes the balance between extremely granular, individual-level analysis that capture too much noise, and overly aggregated information that misses rider habit heterogeneity.

Analysis based on AFC is, however, not without its shortcomings. AFC will miss the travel needs of those who are currently unable to access transit, those who pay by cash, or those boarding without tapping in. This may translate to AFC data failing to accurately capture the ridership patterns of lower-income riders, who are more likely to rely on cash than other transit passenger groups, according to a joint MBTA-Boston Region Metropolitan Planning Organization report in May 2020 [41]. Further, most agencies including the MBTA do not collect key socio-demographic characteristics for those purchasing its fare cards or tickets, so this data is not directly included in AFC. However, card meta-data (e.g., fare type) give us some insight into full-fare versus discount-fare passenger categories, providing a preliminary basis for considering the socio-economic dimensions of each behavioral cluster. Further, by combining AFC data with the latest MBTA System-wide Passenger Survey data at the station-level for rail and route-level for bus, we can give some station- or route-level statistics regarding the socio-demographics of actual riders.

In addition to data analysis, this thesis will contextualize the MBTA's COVID-19 policy response by providing a structured review of how transit agencies in other major U.S. urban areas, as well as cities in other countries, sought to provide public transit services while curbing the virus' spread. The goal of this policy exercise is to collect lessons learned that can be transferable to a Metro Boston context, while also

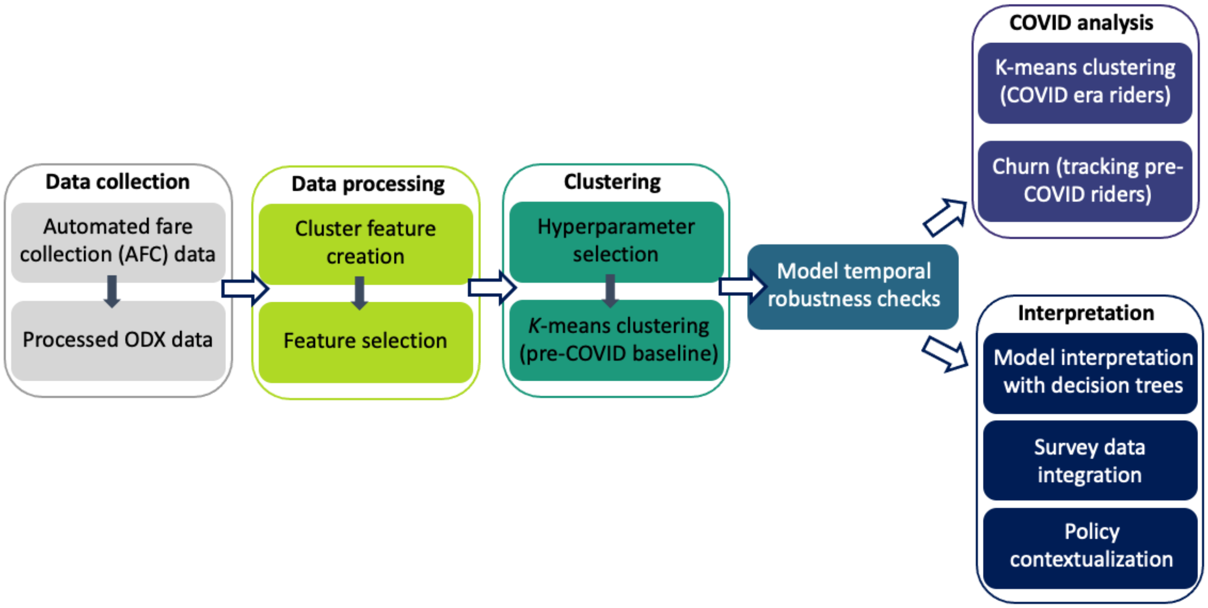


Figure 3-1: Methods flow chart

identifying common pain points in emergency pandemic response that were difficult for most transit agencies to overcome. This policy analysis will mostly focus on pandemic responses, as the majority of urban areas within the U.S. and globally have not yet confirmed full recovery plans as of the writing of this thesis.

The flow chart in Figure 3-1 lays out the analytical flow of this thesis, which is also the flow of this data and methods overview chapter. We begin by discussing our data source, which is an AFC derivative called ODX. We then review common clustering methods seen in the passenger segmentation literature and justify our choice of k-means. Subsequently, we detail the k-means algorithm and the steps we took to pre-process our ODX data in preparation for clustering, the process of creating features from the raw data, and the empirical selection of features for actual use in the clustering step. Next, we describe our selection of hyperparameters including the number of clusters k and the distance measure used to assess similarity of data points during cluster creation. These stages produce our baseline pre-COVID clusters.

Prior to moving on to interpreting the clusters and using them for COVID-19 churn analysis, we need to validate that the k-means model we trained is actually valid across the time periods when we plan to apply it. Therefore, we describe the robustness checks we perform to provide confidence in the model's validity across time. Having validated our model, we then move on to the application of decision trees to interpret the clustering results to give us a more intuitive understanding of how our policy-relevant input features produced the distinct behavioral clusters the model outputted. We also describe our process for giving more socio-economic profiling to each cluster, using both smart card meta-data and integration of a separate MBTA survey. The results of these steps, which are executed on baseline pre-pandemic data and are captured in the blue box in Figure 3-1, will be presented in Chapter 4.

With a clear view of pre-pandemic rider behavior, we then move on to the COVID-19 period analysis. The last two sections of this chapter discuss the application of our k-means model to the pandemic era data, the quantification of churn, and the tracking of passenger behavioral shifts during the pandemic. Finally, we describe the case study "policy matrix" approach we took for overviewing the COVID-19 response of other major U.S. transit agencies, in order to contextualize the MBTA's response. Chapter 5 of this thesis provides the results from the pandemic-era analysis, while Chapter 6 combines the results from our quantitative analysis with our policy matrix work to provide policy recommendations for MBTA's recovery planning.

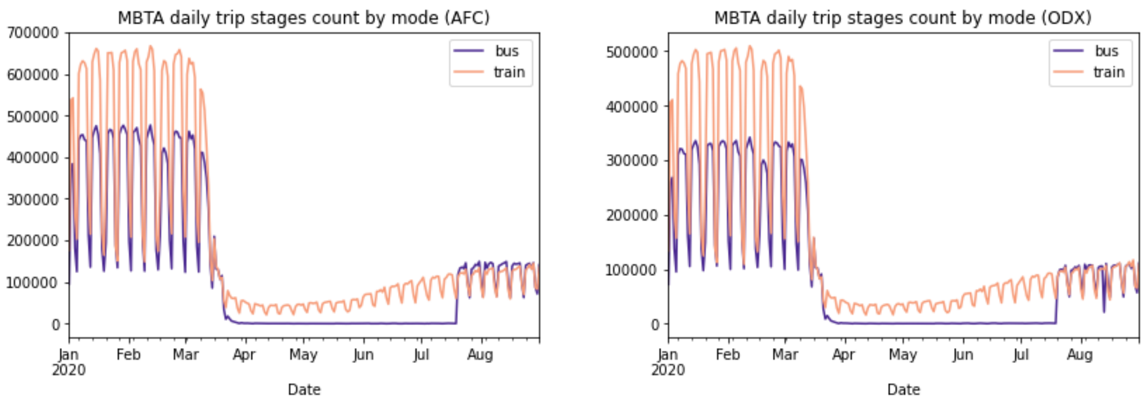
3.2 Data: Automated fare collection (AFC) and ODX

Data for the behavioral clustering analysis was drawn from the MBTA's automated fare collection (AFC) system, which records every smart card or ticket (i.e., CharlieCard or CharlieTicket) transaction. These transactions include taps onto an MBTA vehicle, which is the object of our study. In addition, it includes sales transactions at ticketing machines (e.g., topping up a CharlieCard) and machine errors (e.g., flags for attempted taps into a T station that were mis-processed by the faregate). Because the MBTA hosts commuter rail data separately from the data for its metropolitan "T" subway and bus network, our AFC data does not cover commuter rail.

It is important to note upfront the limitations of the AFC data, since they shape the features we were able to select for our behavioral clustering analysis. Because the MBTA has a flat-fare system where the fare does not scale with the distance traveled, there are no tap-outs required of passengers. AFC therefore only records origins or journey transfers where a tap-in onto a new vehicle is required. This means it does not include information on destinations, and also does not capture in-station transfers where no additional card tap was required for passengers to board their next vehicle (i.e., transfers between subway vehicles). We were therefore unable to take factors such as journey length accurately into account. In addition, because AFC records CharlieCard and CharlieTicket transactions, it groups all cash transactions under one placeholder "card" ID, which covered 1.8% of all taps for our baseline pre-pandemic period of study. Though riders who pay by cash are essential to include in analysis that intends to address transit equity, AFC data does not allow us to do so since it does not provide passenger-level data on cash-paying riders that can be tracked over time.

As noted earlier, AFC contains a variety of transaction data while our analysis only considers the subset representing taps that form a passenger journey. Therefore, we use MBTA's processed Origin-Destination-Transfer (ODX) data as a starting point for our analysis. ODX processes AFC transactions, stripping out records that do not reflect valid taps into the system. It also scales down the AFC data by excluding taps by MBTA system employees. Thus, the volume of transactions shrinks about 20-30% from AFC to ODX, though the temporal patterns of taps are preserved. Figure 3-2 gives a side-by-side comparison of daily recorded trip stages (i.e., tap-ins) on AFC (3-2a) versus ODX (3-2b) through the months of 2020 core to our analysis. As can be seen from the y-axis, ODX recorded fewer trip stages each day, but maintained AFC's modal and temporal patterns. Note also that we lacked data between March 21 and July 20 for bus boardings, due to the MBTA's rear door boarding policy—a pandemic mitigation measure introduced to limit rider contact with the driver. This rear-door policy also affected the surface portions of the Green Line and Mattapan Line trains.

For the work included in this thesis, we focus on the processed ODX data. In order to assess how ridership evolved under the pandemic, we first measured baseline,



(a) AFC

(b) ODX

Figure 3-2: MBTA tap-in data by mode and by source

pre-COVID transit usage. As is evident in Figure 3-2, this meant taking data prior to March 2020. We specifically chose a baseline of January 13 - February 16, 2020, because this was the longest baseline period we could gather while leaving reasonable buffers after the New Year holiday and before we began to see usage declines due to COVID-19 concerns.

This data consisted of 22.6 million trip stages, where each trip stage represents a tap onto an MBTA vehicle and may be part of a longer journey. A record for a trip stage includes information such as the card ID, tap-in time, tap-in location, modal information, tariff information, card meta-data, and the trip stage's position within the journey sequence. We then aggregated the records to the passenger (i.e., CharlieCard or CharlieTicket) level, to allow us to calculate the clustering features discussed in Section 3.4.3 for 1.55 million unique cards, each of which we assume represents a unique passenger. The exact ODX data fields used are listed below.

1. **Card ID:** *Unique ID* of the card/ticket interacting with the system.
2. **Time of tap:** Service *date*, tap *time* (hour, minute, second), *day type* (weekend, weekday, holiday)
3. **Location of tap:** *Location* of AFC data collection. This is the tap-in station for subway rides. For surface transportation including all buses and the Green and Mattapan Line trains which mostly run above ground, AFC only provides

the garage where the farebox is checked, so the actual tap-in location is inferred by ODX using bus schedules.

4. **Trip stage attributes:** How stages fit together into *journeys* as inferred by ODX. *Trip mode* as inferred from attributes of the origin location. *Origin*, ODX *inferred destination*, ODX *route inference* for bus lines based on GTFS. Overall for our baseline period, the 22.6 million trip stages were linked into 19 million journeys.

5. **Tariff and card attributes:** *Ticket stock* allows us to differentiate between paper tickets facing higher per-stage fares, versus cards offering cheaper per-stage fares. An MBTA and Boston Region Metropolitan Planning Organization (MPO) joint technical memorandum in May 2020 on a proposal to eliminate this paper ticket premium showed that low-income and minority riders were historically more likely to use paper tickets (or cash) for the T [41]. *Tariff type* captures whether each stage was taken using a pass validation (for monthly, 1-day, or 7-day passes) or pay-as-you-go. *User group* indicates whether the card/ticket is a regular adult fare product, or one of the special categories. The special categories include blind, TAP (reduced fare for those with disabilities), RIDE (for passengers with disabilities that also book paratransit), senior discounted, student/youth discounted, or short (usually on the bus or Green line, when the vehicle operator allows someone onto the vehicle without paying the full fare). Joining ODX data to a separate database, we can also identify *employer-sponsored* cards. These values will help us profile the socio-demographics of our behavioral clusters. In total, 90% of baseline journeys were taken on adult fare, 5.5% on senior fare, 3.55% on student fare, 0.19% free, and 0.71% short.

We also pulled ODX data for a time period during COVID-19 in order to assess how travel behavior evolved given the pandemic. Because rear-door boarding zeroed out AFC and therefore ODX bus data for late March through late July, we were obliged to pick a period after July 20, 2020 for our pandemic-era analysis. Speaking with MBTA staff and members of local transit advocacy groups, we realized that rear door boarding likely took several weeks to be fully implemented across the system. Therefore, we chose August 18 to September 21, 2020 as our analysis window in order to give the system enough time to normalize. A five-week period was used to match the length of the baseline.

We perform similar five-week data pulls throughout 2019 as well, on which to assess the stability or robustness of our model over neighboring years and over seasons. These validation datasets were pulled for January 14 - February 17, 2019, April 1 - May 5, 2019, August 19 - September 23, 2019, and September 23 - October 27, 2019.

Two key assumptions we make in our analysis of ODX data are 1) each unique smart card or smart ticket ID represents one unique passenger, and 2) each passenger generally uses only using one card or ticket over an extended period of time, so we capture and track his/her full transit usage by monitoring activity on that particular card or ticket. However, these data assumptions may not be strictly true. For example, pay-as-you-go accounts, which are 34% of recorded cards, could allow for instances where multiple riders use the same CharlieCard or CharlieTicket in turn. Additionally, the MBTA is known to have a large share of riders on paper CharlieTickets instead of the more cost-effective CharlieCards. In our baseline period, for example, 56% of riders used CharlieCards and the rest used paper tickets. The paper tickets have shorter shelf lives, leading to concerns that CharlieTicket users could be showing up in the system on multiple days but remain untrackable because they buy a new CharlieTicket and thus appear under a different card ID each time. There is little we can do to directly counter these data limitations, but we keep them in mind as we design our approach and evaluate the results of our analysis.

3.3 Clustering methods

The passenger market segmentation methodology core to this thesis is clustering. Clustering is an unsupervised method for partitioning items in a dataset into clusters, where a high-quality partition groups similar items into the same cluster and keeps clusters distinct. There are many approaches to solving this problem. In the transportation literature on AFC-based rider clustering, the common algorithms used include k-means, hierarchical agglomerative clustering, DBSCAN, and Gaussian mixtures. These are applied to either temporal smart card features, spatial ones, or both to produce a passenger segmentation. Here, we review the common methods and why we adopted the k-means approach using temporal data.

K-means is a widely-used example of a partitional clustering method, in which a group of n points are divided into K groups according to an optimization objective, with the value of k determined *a priori* [42]. Though it requires manual determination of the number of clusters and the distance measure used (often, Euclidean distance), the method is popular because it is computationally efficient on large datasets such as the one used for this thesis, with a run time as low as $O(log\ k)$ when using the k-means++ implementation [42, 43]. However, it only finds local optimums, tend to produce rounder clusters (i.e., it is less effective with correlated features) and is sensitive to outliers [43].

Hierarchical agglomerative clustering (HAC) is a a hierarchical clustering method, which in contrast with partitional methods builds a set of nested clusterings extending from a top "cluster" containing all datapoints down to a set of N final leaves that contains one point per cluster [44]. HAC begins with the datapoints at the bottom and merges the closest clusters until it reaches the single cluster at the top. The resulting tree-like dendrogram is useful for visualizing the relationships between clusters and the distance between them, and a certain number of clusters can be created by cutting the dendrogram at a specified height based on model selection criteria such as BIC or AIC combined with expert judgement [28]. This method's strength lies in this dendrogram visualization, and the fact that it requires little prior knowledge except for a dissimilarity measure. However, finding the point at which to make the dendrogram cut is a well-known problem. Further weaknesses include high time complexity at least $O(n^2 log\ n)$, the process is sensitive to outliers, and merges that occurred earlier in the agglomeration process cannot be undone at a later stage of the process.

DBSCAN is a clustering method that defines clusters by finding areas with greater densities of points than elsewhere in the feature space. Its strength lies in its ability to discover clusters of arbitrary shape, and its robustness to outliers and noise [45]. However, it still relies on the specification of a distance measure and does not perform well if a cluster contains areas of varying density [46].

Generative, model-based approaches learn probabilistic models from the data, with each model representing a specific cluster. They have the advantage of not needing a manual selection of a distance measure [47], as they instead maximize the likelihood of a statistical model fitting the data distribution, which is a more empirically driven approach. Generative clustering models such as Gaussian mixture models follow a similar iterative procedure as k-means. However, instead of assigning a point to the nearest centroid then re-estimating centroids on each iteration, the algorithm assigns a probabilistic membership (soft classification) and re-estimates model parameters based on this. Model-based clustering methods are frequently more interpretable and produce parameter estimates that can be fed subsequently into simulations [46]. However, initialization is still important for avoiding local optima, and the number of clusters k still needs to be manually specified.

As discussed in Chapter 2, hierarchical clustering was adopted by Ghaemi *et al* (2017) and Lathia *et al* (2013) for their analyses. Ma *et al* (2013) and Kieu *et al* (2015) used DBSCAN, while El Mahrsi *et al* (2017) and Briand *et al* (2017) used generative approaches based on mixture models. Briand *et al* (2017), Kieu *et al* (2015), and Ma *et al* (2013) also undertook k-means as a second step in their analyses, while Morency *et al* (2006), Basu (2018), and Fissinger (2020) used only k-means.

This thesis uses a k-means approach. DBSCAN or soft classification using Gaussian Mixture models are more equipped to handle outliers and HAC offers a visual interpretability that could be advantageous for policy [43]. However, they are less efficient to run on large datasets like ours, which contains 1.6 million riders in the baseline and up to 2.02 million riders in the time periods used for model robustness checks [43, 48].

3.4 K-means clustering for customer segmentation

3.4.1 Background on the k-means method

We establish pre-pandemic transit passenger segments by applying k-means clustering to processed AFC data from the pre-pandemic baseline. Given a value for the number of clusters, k, and a data set of n points in R^d, k-means finds a set of k points in R^d to be cluster centroids [43]. That is, using each of the selected features X associated with n observations, we partition the data points into k clusters minimizing the mean squared distance from any point to its closest centroid μ_j:

$$\sum_{i=0}^{n} \min_{\mu_j \in C} \|x_i - \mu_j\|^2 \tag{3.1}$$

We use the k-means implementation in the scikit-learn package written in Python, which approximates k-means with a generalized version of Lloyd's algorithm [49, 50]. The heuristic is based on the property that the optimal placement of a center is at the centroid of its associated cluster. Given any set of k centers Z, for each center z in Z, let $V(z)$ be its neighborhood (ie, its Voronoi cell). Then, Lloyd's algorithm does the following [49]:

1. Initialize Z in R^d
2. Assign each of the n items in the dataset to the closest z, creating $V(z)$.
3. Move every z to the centroid of $V(z)$. Update *V(z)* by recomputing distances from each item to its nearest center.
4. Repeat the above step until the convergence condition is met—i.e., when no data item changes cluster upon recalculation.

This process converges to a local minimum for the sum of squared distances, but is not guaranteed to converge to a global minimum. Effective initialization or seeding of Z to enable more efficient implementations of k-means with higher likelihood of attaining the global optimum has been an extensive field of study [42, 51, 43]. We use the k-means++ seeding method included in scikit-learn, originally proposed by Arthur and Vassilvitskii, to improve our chances of locating the global minimum [42]. K-means++ also has a faster run-time at $O(log\ k)$ [42]. Additionally, we run 300 iterations for each value of k to avoid being trapped in a local solution.

K-means requires a distance measure to capture and minimize dissimilarity. As shown in Equation 3.1, we use Euclidean distance, a common choice in applications like this where features are continuous variables [52, 53, 54]. Typically, clustering based directly on autocorrelated features like time series of AFC taps cannot be done using Euclidean distance and require more complex metrics like cross-correlation distance. However, because our clustering is run on features *extracted* from the time series data, we do not encounter autocorrelation and its resulting complications [55]. Further, Euclidean distance is necessary for the Davies-Bouldin index we rely on to set the hyperparameter, k [56].

Analysis conducted with k-means has to be cognizant of several concerns in addition to the three already mentioned (possibility of falling into a local optimum, dependence on initialization, and computational intensity) [57]. First, this clustering algorithm requires users to manually set k, which is time-consuming and subjective in unsupervised cases where the natural groupings in the data are not known. To address this, we take the commonly used approach of leveraging the Davies-Bouldin index and elbow heuristic to give scientific guidance to our hyperparameter selection (Section 3.4.4). Second, k-means is sensitive to isolated outliers which can pull calculated centroids far from the optimum. We address this by removing clear outliers from our data prior to clustering (Section 3.4.3). Third, k-means tends to produce rounder and more evenly sized clusters. This relates to the development of k-means as a clustering method based on variance without consideration of covariance, which also means we must avoid using highly correlated features [43]. We discuss our feature selection process in Section 3.4.3.

Despite its limitations, k-means is extensively used on large-scale human activities data because of its relative efficiency in the unsupervised clustering setting. Extensive work has been done to enable faster, more efficient implementation of k-means over larger datasets, without requiring unreasonable data storage capacity [58, 59, 43]. Dong et al. (2015), for instance, applied this method to call detail records in Beijing for the purpose of partitioning the city into land use zones and traffic zones. Zhang et al. (2020) used k-means on transit smart card data to identify areas of Wuhu, China where vulnerable populations follow similar travel behaviors—a setting similar to the topic of this thesis.

3.4.2 Data pre-processing

As a first step, passengers were separated into two groups based on the number of days they were active in the system. Those with one or more trips on only a single day were put into the single-day category, while those with trips on multiple days were put into the multi-day category. We made this high-level separation prior to creating features for k-means. Since we cluster on temporal attributes, many potentially useful measures of ridership intensity—e.g., days of active usage—take on only a single value (i.e., one day) for the single-day users that make up 43% of cards recorded in the pre-COVID baseline. Thus while these ridership intensity features have variability and can be key to partitioning multi-day riders, they do not contribute to the large portion of single-day riders in our data, who further can reasonably be expected to behave in ways distinct from recurrent users. Lastly, the transient nature of the single-day riders' system engagement is highlighted by the fact that 75.5% of single-day users possessed the more temporary CharlieTickets; by contrast CharlieTickets were only 19.8% of multi-day users during the baseline. This type of manual separation between major rider groups has been used in previous literature where clear behavioral distinctions were known *a priori* [46, 32].

There is potential for mis-categorization for those who entered or left the system at the boundaries of our sampled timeframe—e.g., multi-day users who stopped riding the MBTA on January 13 would be miscategorized as a single-day user even if they had ridden every day in the preceding week. However, we assume that this effect is small compared to our full dataset. As we will see in chapter 4, the clear-cut nature of our single-day rider results suggest there was not much contamination from multi-day users mis-categorized due to "boundary effects." Overall, we found that 57% of card IDs were multi-day, and the remaining 43% single-day.

3.4.3 Feature creation/selection

With processed datasets in hand, the next step is to identify features within the data with which k-means can efficiently partition the passengers. Some smart card-based studies apply clustering techniques directly to the raw time series of taps. For example, He, Agard, and Trepanier (2017) used this type of approach with hierarchical clustering, segmenting transit riders in Gatineau, Canada based on daily ridership behaviour [55]. Briand et al (2017) applied a Gaussian mixture generative model to disaggregated time series of rider taps [14].

Meanwhile, other authors constructed features from the raw data prior to analysis. El Mahrsi *et al* (2017) aggregated taps into 1-hour bins [46]. Basu (2018) applied principal component analysis (PCA) to AFC to create orthogonal features that work well for k-means [26]. Zhang et al (2021) calculated more aggregated passenger-level metrics such as travel frequency and time entropy to use as features [54]. Ma *et al* (2013) created features from temporal data, like number of active days logged by a smart card and indicators for similar boarding times [29].

The diversity of methods for feature creation in the literature reflects the multiple approaches to this question of dimension reduction. The approaches usually fall into two categories. One is feature selection, which searches for irrelevant original features to exclude, and also sometimes applies feature weighting to selected features. The other is feature extraction, in which new features are created from the originals. This thesis, similar to Fissinger (2020), Ma *et al* (2013) and Zhang *et al* (2021), first uses feature extraction to create key temporal variables from ODX data; it then uses feature selection to pick out the best extracted features for k-means [32, 29, 54]. Since our study aims to provide interpretable input features to the clustering that can be directly translated into urban planning dialogue, we chose this approach of extracting more policy-relevant features from the ODX time series for clustering, rather than attempting classification on the raw time series.

We also decide to apply k-means only to temporal data, while as discussed in Chapter 2 the literature sometimes clusters based on spatial data as well. Temporal patterns are useful for understanding user habits and network flows, which are essential for system operations and tracking how passenger habits evolved during COVID-19. Spatial considerations are also essential for transit planning and understanding the pandemic's equity impacts on transit usage. The temporal dimension is, however, the one that is ubiquitous across passenger segmentation studies, with multiple authors finding that the frequency and temporal profile of daily ridership are especially key for segmentation [14, 46, 28, 31, 30]. Further, working with spatial data for the MBTA is limited by the fact that AFC does not collect alightings, only boardings. Using spatial features would also require inferring home locations and assuming some similarity of riders to the neighborhoods of their trip boardings [32]. This would add further uncertainty to the data that forms the basis of our clustering. Thus, we conservatively use the spatial information from AFC only during our profiling stage.

Our feature extraction process closely follows Fissinger (2020) and Basu's (2018) procedure prior to the PCA. We focus on creating temporal variables that capture the frequency and timing of rides, and variables that reflect modality which is central to operations planning [26, 32]. Final features design was done with input from the MBTA. Basu's subsequent PCA step has the advantage of producing an orthogonal basis supportive of more robust k-means clustering. However, PCA results are generally of low interpretability and therefore less fitting for this thesis's audience [60]. We generate the following set of possible features from among which to select:

1. **% peak journeys:** The share of journeys on each card/ticket that begins during peak hours (6:00-9:00 and 15:30-18:30). While MBTA morning peak hours are officially 6:30-9:00 AM weekdays (i.e., that is when service runs more frequently), our partners at the MBTA advised we start our morning peak window at 6:00 AM, since a notable share of commuters have embarked on their travels by then. The daily transit ridership profile in Figure 5-3 confirms this. Meeting peak demand is critical for transit agency operations, and reflects

one of the biggest areas of ridership change during COVID-19. Overall, 64% of cards in the baseline had peak-period journeys.

2. **% journeys with a transfer:** The share of trips on each card/ticket with a bus-related transfer, as labeled through trip chaining by ODX. In MBTA data, rail-to-rail transfers are not captured as they are within-station, and ODX inference accuracy for it is unverified. Further, from a rider experience perspective, transferring outside for a bus is a more significant and negative experience than transferring trains inside a station, so there is some justification for differentiating between rail-rail and bus-rail transfer experiences. Only 24% of all cards in our baseline period had bus-related transfers.

3. **% weekend trips:** Share of trips on each card/ticket taken on a weekend or holiday, as defined by the MBTA weekends and holiday service calendar. Overall, 42% of cards had weekend or holiday travel, and the distribution is fairly bimodal with modes at either end of the range.

4. **% trip stages by bus:** Share of trip stages on each card/ticket taken by bus. Overall, 41% of smart cards had bus-based stages in the baseline.

5. **Active range (days):** The number of days between the first and last use of the smart card during the study period. This is used as a measure of ridership persistence. In the baseline, the peaks of the distribution are at 1 day and around 29+ days (out of 35 total).

6. **Average trips per active week:** Average number of trips taken during an active week (i.e., excluding weeks with no trips), used as a measure of usage intensity.

7. **Active days:** Number of days during the study period when the smart card was used.

8. **Active weeks:** Number of weeks during the study period when the smart card was used. This is highly correlated with active days, and we anticipated one of

these will drop out during the feature selection process.

The last four potential features were normalized to [0, 1] to align with the magnitudes of the percentage variables, since variations in cardinality can interfere with k-means.

We then picked our final feature sets for multi-day users and single-day users, separately. The goal was to select a subset of possible features that is useful for customer segmentation, but also sparse enough to be interpretable. We use two common feature selection methods, greedy sequential forward selection and sequential backward selection [60]. Greedy selection starts from an empty feature set, and on every step adds the most rewarding feature from the unselected ones. Because this tends to over-select features, we compare this against backward selection, which starts from the full feature set and removes the least rewarding feature each round until some stopping criterion is hit. In either case, the decision of whether a feature is "rewarding" is typically made by ranking using entropy scores, redundancy measures (which penalizes highly correlated features), or relevance measures (an opposite approach favoring correlated features, assuming correlation reflects cluster structure) [60, 61].

Our method is closest to ranking by entropy scores. We conduct a grid search running greedy and backward selection for $k = 5$ to 7 for multi-day users and $k = 2$ to 5 for single-day users, randomizing the order of the features since these sequential selection methods are influenced by the order of the features tried. Each time a feature is added or removed, we calculate the "correspondence score," defined as the share of datapoints that did not shift clusters between different model runs [26, 62].

Specifically, we take the points in the baseline period and cluster it using the existing feature set to produce labels L_1. We then cluster it again using the new proposed feature set to produce labels L_2. Thus, each passenger now has two labels, generated by models trained on the two different feature sets. If arranged in a matrix with L_1 as the rows (i) and L_2 as the columns (j), two models that produce similar clustering results would therefore have high values along the diagonal (i.e., more matched labels represented by the elements a_{ii}) and low values away from the diagonal (i.e., fewer mis-matched labels represented by the elements a_{ij}). Using this structure, the correspondence score is defined as follows:

$$correspondence\ score = \frac{\sum_{i=1}^{i=n} a_{ii}}{\sum_{j=1}^{j=n} \sum_{i=1}^{j=n} a_{ij}} \tag{3.2}$$

This definition allows us to quantify the level of "similarity" or "agreement" between two models applied to the same dataset. If the correspondence score is above 90%, it means that over 90% of the data points did *not* change labels with the addition/removal of the latest feature under consideration. At this point, we consider the latest feature assessed to be non-essential and discard it.

Though we also noted the changes in entropy during the feature selection process, we did not use entropy scores directly for our evaluation. This is because the meaning of any particular entropy value is hard to interpret, and therefore it is difficult to gauge where to set the appropriate cut-off for variable inclusion. Meanwhile, correspondence scores give a clearly interpretable way to assess the clustering impact of the latest feature under consideration. The features selected by this process for the multi-day and single-day user groups will be presented in chapter 4.

Finally, as previously noted, k-means is sensitive to outliers. We therefore excluded passengers with clear outlier values along selected features. This translated to excluding 1,549 riders with average journeys per week in the top 0.1% (i.e., over 25.3 journeys/week).

3.4.4 Selecting the hyperparameter k

In k-means clustering, we must provide a value for the hyperparameter, k. Common methods for finding k include the elbow heuristic [54], Davies-Bouldin Index (DBI) [56], and silhouette scores [60]. We choose DBI and the elbow heuristic because these methods are computationally easier to calculate than silhouette scores for larger datasets. DBI aims to minimize intra-cluster variance and maximize inter-cluster distance, thereby identifying distinct, compact clusters. It is defined as the average similarity between each cluster C_i and the one most similar to it, C_j. The similarity metric R_{ij} trades off the cluster diameter s_i and the distance between the cluster centroids of C_i and C_j. We use the sklearn implementation of DBI in Python, which defines R_{ij} as

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \tag{3.3}$$

So DBI is:

$$DBI = \frac{1}{k} \sum_{i=1}^{k} \max_{i \neq j} R_{ij} \tag{3.4}$$

DBI only requires calculating distances of each point from its nearest centroid; this is much less intensive than silhouette scores, which requires calculating distances between every pair of points. Thus, DBI is more tractable for large datasets like ours. The elbow method, on the other hand, is even simpler–it plots the fall in inertia[1] as k increases, and the range where the rate of decrease in entropy slows as more clusters are added marks the optimal number of clusters. Further, DBI uses Euclidean distance, which works well with our data's continuous features. DBI also requires features to be normalized or standardized, since wide variations in the cardinality of features would skew outputted values [60]. This is another reason behind our decision to normalization features to [0,1].

There are additional improvements that can be made to our methodology. Fsea-tures can be normalized by cross-projection to reduce bias [43]. Silhouette scores, which has been validated as performing better than DBI in selecting k, could be run if more computational power were available. Other clustering methods could also be used and their outcomes compared, given greater computational power to handle the longer running times. For example, DBSCAN or Gaussian mixtures can be leveraged to diminish the influence of outliers [14].

[1] The sum of squared distances.

3.5 Robustness check: Temporal stability of customer segments

As discussed in Section 3.2, the baseline pre-COVID period upon which we base our analysis is a 5-week period from January 13 to February 16. However, this baseline period is in the winter and is also reflective of conditions in early 2020. There is the possibility that in other seasons and in other years, ridership patterns would point to a different customer segmentation; i.e., there is natural temporal drift in customer behavior and segmentation. Since the COVID period stretches across multiple seasons and at least two years, having a stable baseline customer segmentation is essential for not confounding ordinary background shifts in ridership patterns, such as those studied in Briand (2017), with COVID-induced shifts [14]. Therefore, we need to perform robustness checks for our k-means model against data from at least one other year and non-winter seasons.

We perform these checks with 5-week periods in each season of 2019, similar to the procedure in Goulet-Langlois (2016) [62]. This includes January 14-February 17 2019 for the winter, April 1-May 5 2019 for the spring, August 19-September 23 2019 for the summer, and September 23-October 27 2019 for the fall. The summer sample is close to the fall sample because we chose it to match our 2020 COVID period, which was late in the summer due to the fact that bus AFC data was zeroed out until July 16, 2020 because of MBTA's bus rear-door boarding procedures for COVID-19 prevention.

For each robustness check period, we followed the same process as previously described to train a k-means model on that sample and discover underlying behavioral clusters. Then, we separately run each sample through the k-means model trained on the baseline period data. This means we generated two separate cluster partitions for each stability check period—one using the model trained on the baseline data, and another using the model trained on the robustness check period data itself.

We then apply correspondence scores to compare the results from the two models, as described in equation 3.2. Note that unlike before, the correspondence score is being used to compare models derived from two different sets of training data, instead of two different feature sets drawn from the same training data. Again, a higher correspondence score indicates greater similarity or agreement in the classification decisions of the two models, and therefore greater temporal stability of models trained on two different time periods.

To calculate appropriate correspondence scores, we also need to identify which clusters from one model counts as a "match" with a cluster from the second model. We do so by assuming that the label from the second model most frequently represented in the members of a particular label produced by the first model is the matching cluster. For example, if Cluster A from Model 1 is labeled as Cluster D 95% of the time by Model 2, and only 5% of the time as Cluster A, B, or C, we will assert that Model 1 Cluster A is matched to Model 2 Cluster D.

Even though we can automatically identify matching clusters between pairs of models using the above procedure, it is still worth supplementing the correspondence scores through a qualitative comparison of the centroids. This allows us to assess how similar the representative point in each cluster is across models, when measured against the clustering features we have chosen. Finally, we examine the percentage of total riders in each cluster produced by the robustness check versus the baseline period model, to see whether the *distribution* of cluster membership has shifted.

3.6 Optimal Classification Trees (OCT)

The goal of this thesis is to create and track interpretable, policy-relevant customer segments through COVID-19 in Metro Boston. Therefore, we examine optimal classification trees (OCT) as a method for offering a more intuitive understanding of the key factors that partition riders into different behavioral groups. The OCT method used is based on Bertsimas, Orfanoudaki and Wiberg (2019) and Bertsimas and Dunn (2017), with an implementation the authors made freely available online at https://www.interpretable.ai/ [63, 52].

Decision trees are widely used for classification, and are often favored for their human-interpretable outputs compared to more accurate but much less parsable methods [63]. However, one disadvantage of single decision trees (as opposed to ensembles like random forests, which are again less interpretable) is that the splits in the tree are chosen in isolation, without taking into account the forward implications of the future splits a particular decision generates. OCT skirts this problem by applying mixed-integer optimization (MIO) to form the entire decision tree in a single step, thereby allowing each split to be determined with full knowledge of all other splits. Because it considers forward splits, it is not myopic like typical decision trees and can better capture the structure of the underlying dataset. To avail ourselves of the legibility and interpretability of OCT in an unsupervised clustering setting, we use the labels produced by k-means to turn the problem into a supervised one and apply an optimal decision tree (OCT) to observe the decision criteria upon which the clusters were originally split by k-means [52].

Building on the classification and regression tree (CART), the univariate OCT splits along one decision variable at a time. However, because the problem is solved in one step, there is no longer a need to update impurity measures when growing the tree as is done in more traditional decision trees. The "optimal tree problem" formulation given by Bertsimas and Dunn (2017) is reproduced below. $R_{xy}(T)$ is tree T's misclassification error on the training data (x_i, y_i) for i = 1,...,n. A complexity parameter α is used to control the complexity-accuracy trade-off for T. |T| is the count of T's branch nodes (hence the penalty term is attached to it). The number of points in leaf node l (denoted $N_x(l)$) must be at least N_{min}.

$$\min \quad R_{xy}(T) + \alpha|T| \tag{3.5a}$$

$$\text{s.t.} \quad N_x(l) \geq N_{min}, \tag{3.5b}$$

$$\forall l \in \text{leaves}(T) \tag{3.5c}$$

The hyperparameters are complexity and tree depth. These were tuned using grid search; a smaller complexity penalty and greater depth leads to a "shrubbier" tree that produces fewer misclassifications but is more complex to interpret.

This method for using OCT to aid in the interpretation of clustering problems is drawn from Bertsimas *et al* (2019). However, due to computational limitations and the large size of our dataset, we do not follow the recommended approach of simultaneous deciding the clustering and the OCT structure. Further, the authors of the referenced paper have not yet made packages implementing the simultaneous method available. Instead, we use the alternative method in which we first run k-means to generate labels for our baseline dataset, and subsequently apply OCT. Though this approach does not produce optimal feature selection and clusters like the simultaneous method, this is acceptable in our use case since our purpose is to interpret rather than to further optimize our clusters. The result of this portion of our analysis is one OCT for the multi-day passengers, and a separate OCT for the single-day passengers.

3.7 Cluster profiling

We cluster on the temporal and modal attributes captured by smart cards, that is, on rider behaviour only. However, MBTA AFC data also provides smart card meta-data such as the use of passes versus tickets, company sponsorship, and card type (e.g., discounted cards for vulnerable populations). Thus, to help us understand the demographics of each cluster and equity implications across clusters, we cross-tabulate clusters against the available smart card meta-data. The meta-data is available for all cards in use at any time, so we are also able to assess how the profiles changed from the baseline to the COVID-19 period.

ODX also includes boarding locations for each trip stage. This allows us to conduct some spatial profiling of riders in each cluster, by 1) examining the top 10 boarding locations listed for each cluster, and 2) assessing the entropy of boarding locations at the passenger level and comparing the distributions of boarding location entropy across clusters [14]. To measure this, we calculate the Shannon entropy for each passenger, with i indicating a station or bus stop [64]:

$$H(X) = -E[logP(X = x_i)] = -\Sigma_{i=1}^{n} P_i log P_i \tag{3.6}$$

Entropy represents the probability of a smart card validation at stations most active for the card, and is a measure of the dispersion/variation in station usage for that card. This gives us a sense of the geographical breadth across which members of each cluster tend to use transit, with higher entropy scores corresponding to greater geographical spread.

For the purposes of cluster profiling, we also complement our AFC/ODX data with data from periodic customer surveys conducted by the MBTA. The latest 2015-2017 System-wide Passenger Survey captures rider socio-demographics, automobile availability, and use of alternative modes. The data was collected by surveying riders at each subway station and on board buses along each bus route [65]. This can be matched to our trip origin and bus route data to further profile the rider composition of each cluster in policy-relevant ways.

However, no similar survey was conducted during COVID-19. Thus, unlike the smart card meta-data which is available for all time periods, the survey can only be used to give us a sense of the baseline demographics and car ownership patterns common across clusters. It cannot be used for tracking how these profile characteristics evolved during COVID-19. To do so with high granularity would require an additional survey to be conducted during the pandemic.

Differences in the method of data collection for the survey versus AFC/ODX also limit our ability to fuse these datasets for analysis. For its rail component, the passenger survey was conducted by soliciting anyone in a subway station, whether they were entering, transferring or exiting—this is in contrast to AFC/ODX, which captures tap-ins (i.e., only those entering a station). Therefore, joining survey data to our ODX database by station is not completely accurate. For bus and above-ground Green Line stations, the survey data is mostly at the route-level. This gives us very limited specificity, especially considering that many bus and Green Line routes pass through neighborhoods with diverging socio-demographic characteristics. Given these constraints, we only use the survey data to get a broad sense of where differences in baseline cluster socio-demographics may be, and how observed travel behaviors are correlated with factors such as car ownership and other available modal alternatives.

3.8 Churn analysis during the COVID-19 pandemic

There is limited literature tracking the evolution over time of customer segments, especially during extended abnormal emergency periods like COVID-19. Our work in this domain most closely follows that of Briand et al (2017), which tracked the composition of clusters and churn rates over five normal time periods [14]. However, unlike Briand et al (2017), we specifically wish to describe the *incremental* churn associated with the *abnormal* COVID-19 event. We thus need to compare the atypical situation of tracking clusters from baseline into COVID-19 periods, against the "background" or "natural" churn seen during similar seasons in normal years. Therefore, in all our churn analysis, we compare the cluster evolution from baseline 2020 (winter) to pandemic-period 2020 (summer) against winter to summer 2019.

We examine change over time in two ways:

1. **Examine all rider accounts during COVID-19, including both new and pre-existing passengers.** We apply the model trained on the baseline period data to the full set of pandemic period users to understand how cluster membership and the share of each cluster in the overall population changed during COVID-19. This gives us a comprehensive view of ridership patterns during the pandemic, including those who decided to enter the MBTA system during the crisis. This view is helpful for the MBTA as a transit operator trying to better understand the full picture of the new demand patterns they are facing during the pandemic.

2. **Track pre-existing baseline-period riders through COVID-19, to assess churn and cluster-switching.** This gives us a narrower, longitudinal view of how riders who were with the system at the start of the year chose to

avoid transit, or shift their pattern of usage, in response to the pandemic and associated restrictions and transit service changes. To produce this analysis, we still apply the model trained on the baseline period data to the summer 2020 data, but only for cards that had already registered in the system in January/February. While this approach gives a less complete view of pandemic-era travel behavior, it gives us a clearer sense of how individuals in the system adapted to the crisis.

We apply the same two methods to compare winter 2019 versus summer 2019 data, in order to create a benchmark for "background" churn and cluster switching. Further, we again utilize smart card meta-data to profile clusters during COVID-19, in order to determine whether stickiness to transit usage is unequally distributed across socio-demographic segments. Lastly, we calculate Shannon entropy and take stock of the most active stations, to get a sense of whether the geographic spread of transit usage and the busiest locations in the system have changed due to COVID-19.

An additional method that could be applied for pandemic-period analysis is to re-train a k-means model on the COVID-19 period data to generate a new clustering, which can capture any shifts in cluster number or centroids during the pandemic. Because the clusters produced from this may be defined differently than the ones from the baseline model, this approach makes it more difficult to compare baseline to COVID-19 era for measuring churn and cluster-switching in an interpretable way. However, future work providing a pandemic-period clustering and delving into methods for quantifying shifts in cluster centroids could provide a valuable complementary analysis on how COVID-19 shifted the landscape of ridership behavior.

3.9 Policy Matrix

The quantitative analysis provided a window into the demand side of pandemic-era transit usage. However, it does not capture the supply side, or the stakeholder and planning complexities facing transit agencies trying to adapt to an unprecedented situation unleashing widespread uncertainty. To contextualize the policy-relevant conclusions drawn from our quantitative analysis, we studied COVID-19 transit agency responses across Boston and 12 other U.S. urban centers with relatively extensive public transit systems. We also assess 12 international urban centers in Asia, South America, and Europe, to broaden the number of case studies from which we can draw lessons learned.

For each system, we collect the following information in order to get a view of the natural state of the agency, how it was impacted by the pandemic, and how it responded.

1. **Transit usage and system design** prior to the pandemic. Differences in system design across cities can strongly influence the pandemic's ridership impact and the transit agencies' response. We take this into account in order to better assess where cities are comparable, and where they are not.

2. **Transit agency structure and funding mechanisms** during normal years. The structure of authority and the financial health of transit agencies prior to the pandemic can influence the rapidity of the agencies' responses, and the amount of funding they can dedicate to sanitation efforts, maintaining service frequency, and ramping up for recovery.

3. **COVID-19 severity and its impact on transit ridership and revenue**, including any pandemic-related funding received. Here, we seek to clarify ways in which the pandemic had similar impacts on other major transit agencies, and areas where its impact may have differed due to variations in COVID-19 response. Revenue and emergency funding are also key variables to track if we are to understand the constraints upon agencies' ability to effectively operate during the pandemic.

4. **COVID-19 responses** by transit agencies, including mask mandates, service cuts, sanitation, ventilation, social distancing guidelines, case tracking, COVID-19 communication, and enforcement of instituted prevention measures.

We synthesize this information into a matrix format to facilitate comparison across cities (matrix columns) for each data category (rows). The main goal of this comparative exercise is to understand whether the MBTA's responses aligned with that of other transit agencies, whether there are lessons learned from other locations that can translate into the MBTA context, and whether there are common obstacles that posed an insurmountable challenge to all agencies. We also highlight structural as well as cultural differences that can limit transfer of policies across cities.

Chapter 4

Baseline transit rider clusters: Results and interpretation

4.1 Overview

Applying the k-means clustering methodology to ODX data in our baseline pre-pandemic period of winter 2020 produced seven multi-day and three single-day MBTA rider clusters. The two largest multi-day clusters represent categories of bus-oriented riders, while the remaining five capture variegated ridership behavioral among rail-oriented riders, with occasional rail riders providing the greatest variety. Among multi-day clusters, the results of optimal classification trees (OCT) indicate that passengers' primary transit mode (bus or rail) form the first major variable upon which they were partitioned in the process of sorting passengers into behavioral clusters. This is followed in the tree structure by the frequency of transit use and finally the particular time of day during which a passenger typically enters the transit system. Single-day riders, on the other hand, were rail-dominated and fell into clean categories based upon the time of day and time of week of their ridership.

Figure 4-1 summarizes the features selected for clustering, and the multi-day as well as single-day clusters that resulted from the k-means procedure. Multi-day clusters are numbered 1 through 7, and single-day 8 through 10. Separately within the multi-day and the single-day sections, clusters are listed in descending order of passenger count. The values listed in the leftmost panel of the table are generally k-means cluster centroids. However, unlike for the multi-day, for the single-day only peak share and weekend share were selected as clustering features. Thus the values provided for the remaining features of single-day clusters are averages rather than k-means centroids.

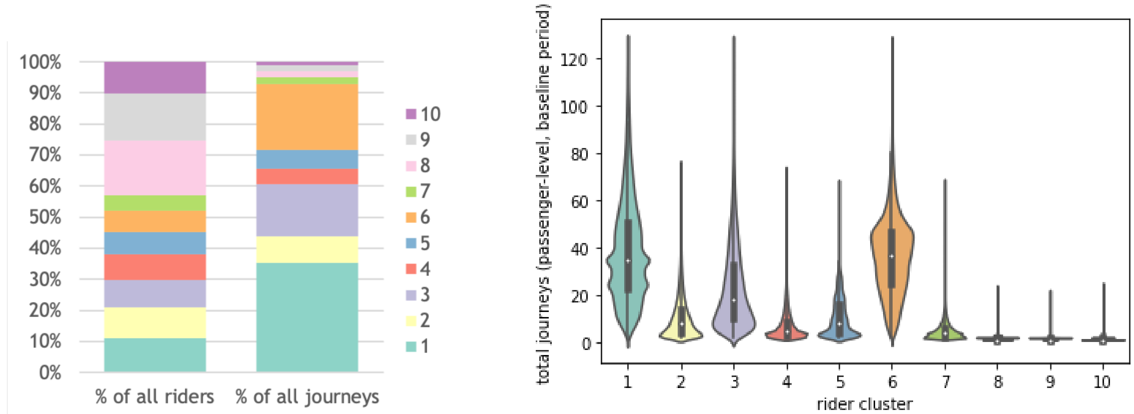
Cluster #	Cluster name	bus share	peak share	active range	weeken d share	transfer share	active days	% of all riders	% of all journeys	# of riders	# of journeys
1	Frequent bus riders	75%	45%	32.0	14%	27%	20.1	11%	35%	172,552	6,715,417
2	Occasional bus riders	76%	37%	11.9	14%	26%	5.5	10%	8%	152,429	1,599,178
3	Fairly frequent wkday off-peak rail	11%	24%	29.3	18%	4%	12.9	9%	17%	134,790	3,201,491
4	Occasional wkday off-peak rail	6%	25%	9.2	7%	3%	4.1	8%	5%	127,368	918,451
5	Occasional rail commuters	8%	85%	11.5	2%	4%	6.1	7%	6%	110,398	1,171,138
6	Frequent rail commuters	8%	74%	31.9	5%	5%	19.4	7%	21%	109,225	4,030,200
7	Occasional weekend riders	15%	10%	11.1	68%	4%	3.2	5%	2%	73,976	391,277
8	Single-day, off-peak weekday	14%	1%	1.0	0%	3%	1.0	18%	2%	271,676	369,198
9	Single-day, peak	12%	85%	1.0	0%	3%	1.0	15%	2%	236,792	348,397
10	Single-day, weekend	11%	0%	1.0	100%	2%	1.0	10%	1%	155,101	225,299

Figure 4-1: Feature centroids by cluster for multi-day users and single-day users, cluster share of total ridership and total journey count in baseline 2020

Figure 4-2a and 4-2 illustrate the relative size and thus operational importance of each cluster, in terms of their share of the total rider pool and their contributions to the total journey count. From these charts and the heatmap table, it is evident that the large number of single-day riders are extremely small contributors to the actual usage of the MBTA network. Thus, we focus most of our analytic attention in this and upcoming chapters on the multi-day rider clusters.

The rest of this chapter fleshes out the cluster results summarized above, then provides interpretations of each behavioral cluster based on complementary data pulled in to give a clearer profile of the passengers sorted into each cluster. Following this structure, we start with the results of cluster feature selection, k selection, and k-means results—for multi-day, then single-day clusters. We then report on our temporal robustness checks, which indicate that the baseline k-means model we trained on winter 2020 pre-pandemic data is applicable across adjacent years and between seasons. This gives us more confidence going into the subsequent COVID-19 analysis chapter, in which we must apply our baseline model trained on winter 2020 data to ridership behavior during a different season, summer 2020.

Next, this chapter profiles the passenger composition of each cluster using two



(a) Cluster contributions to passenger count and journey count

(b) Distribution of passenger journeys by cluster, baseline period

Figure 4-2: Cluster size in terms of passengers, journeys

data sources—smart card meta-data included in AFC and ODX, and less granular data from the MBTA's 2015-2017 System-wide Passenger Survey that gives information on demographics and alternative modes. The smart card meta-data captures each passengers' usage of full adult fare products, reduced or free fare products, and corporate-sponsored Perq passes. Analysis suggests that frequent bus riders, occasional bus riders, and fairly frequent weekday off-peak rail riders are the behavioral clusters capturing the majority of reduced-fare passengers, though they differ in the sub-segments of reduced-fare riders that they tend to represent. The System-wide Passenger Survey complements these conclusions, suggesting that the bus clusters in particular may capture lower-income households, people of color, and those with fewer alternatives to transit use. These conclusions suggest the evolution of these bus clusters and the fairly frequent weekday off-peak rail rider cluster needs to be carefully monitored going into the pandemic and post-pandemic recovery, to support service for those whose behavior indicate greater levels of transit dependence.

Perq smart card meta-data indicates frequent rail commuters make up a disproportionately large share of all Perq users (39%) compared to their share of the general rider pool (7%). Since Perq historically contributes one-third of MBTA fare revenues which are in turn around one-third of operating revenues, this suggests that the frequent rail commuters cluster is key to post-pandemic revenue recovery [66]. However,

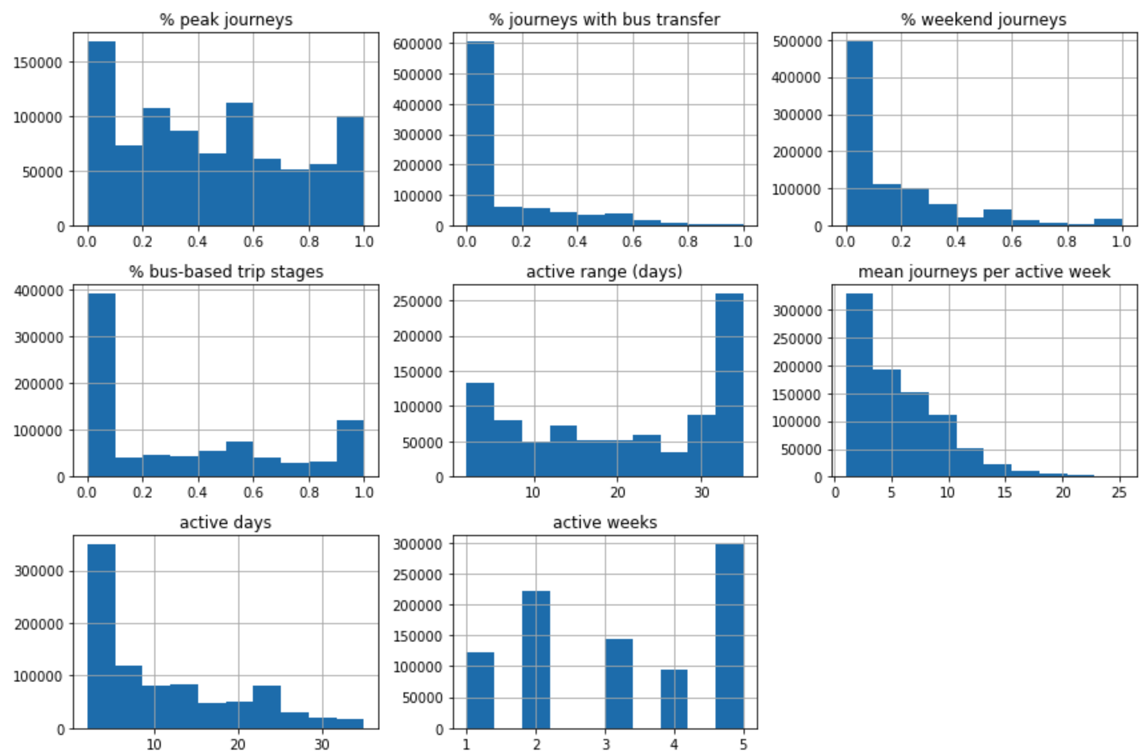


Figure 4-3: Distribution of candidate clustering features for multi-day riders

managing ridership and revenue recovery by this group will look very different from managing recovery for other clusters, because it will be heavily influenced by MBTA negotiations with corporations subscribing to Perq.

4.2 Baseline transit rider clusters: Multi-day riders

4.2.1 Selected clustering features

As discussed in Chapter 3, our clustering analysis was performed separately for the multi-day and single-day riders because they offer different sets of temporal features as candidates for clustering. There were 880,738 multi-day riders in our baseline data, or 57% of all riders that interacted with the system during that time. Of these, 80.2% were CharlieCard users, rather than users of the more temporary paper CharlieTickets—as could be reasonably expected of recurrent transit users.

We began the feature selection process by considering all eight features created from the underlying ODX data. Descriptive analysis like Figure 4-3 shows that, because most passengers had many trips over a substantial number of days, the multi-day passenger data is much smoother compared to the single-day passenger data that we will later show in Section 4.3, although there are still longer right tails especially for the share of journeys with bus transfers. The share of peak journeys taken by commuters is relatively evenly distributed, indicating a mix of commuters and non-commuters in the data. Meanwhile, the share of weekend journeys and mean journeys per active week show substantial variation as well, but mostly in the lower half of the variables' ranges. The share of bus-based trip stages and the active range were relatively bi-modal, suggesting the dataset may be split between bus-heavy users versus rail-heavy users, and between frequent versus more occasional riders. Further, the contrast between the bi-modal nature of active range (and also active weeks), versus the right-skewed shape of active days, suggest that there is a substantial group of riders who used transit over a large time window, but were not using the system at a very high frequency within that window (e.g., riders who rode twice per week for all five weeks, as opposed to riders who rode five times per week for all five weeks).

We also check the correlation matrix between the candidate features. Since k-means does not capture covariance between features, we prefer the final feature set to show low correlation. Figure 4-4 displays the correlation matrix, and we examine this with the heuristic that correlations greater than 0.7 in absolute value indicate pairs of features that should not jointly enter our final feature set. The figure shows that active range is, as expected, highly correlated with active days (r = 0.75); active range is also highly correlated with active weeks (r = 0.93). Active days and active weeks are highly correlated at r = 0.8. We therefore expect that only a subset of less correlated usage intensity features will be selected as our final feature set.

On the other hand, the correlation between peak period ridership and weekend ridership are only of medium strength at r = -0.47, because there is substantial off-peak weekday ridership in the system. Further, the correlation between passengers' share of trip stages by bus and share of journeys with transfers is fairly limited at 0.39, reflecting the large share of non-transfer trips observed in Figure 4-3. Even if

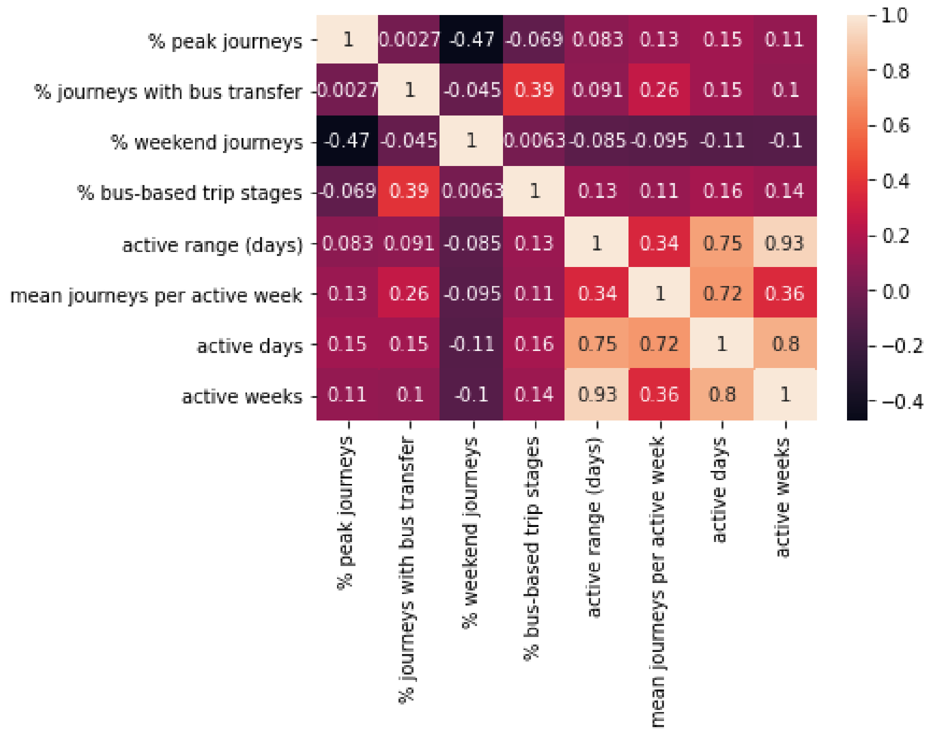


Figure 4-4: Correlation matrix among candidate clustering features, multi-day riders

we filter out the 39.6% of passengers who did not use bus at all during the baseline period, we find that mean bus-related transfer share is relatively low at 20.2%, though the longer right tail remains. This alleviates our concern that bus share and transfer share would be heavily correlated due to the nature of AFC data, which only captures bus-related transfers.

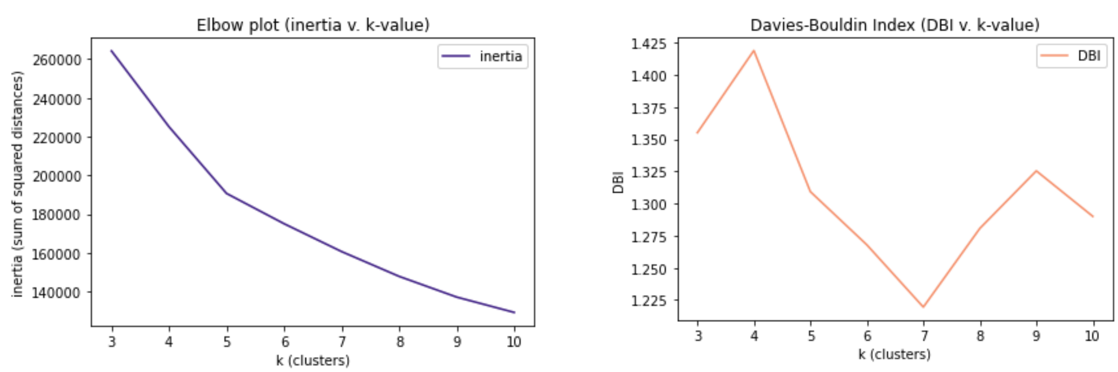
We ran the greedy and backward feature selection methods for k values of 5 through 7, since an initial elbow plot analysis with all candidate features suggested the optimal k may be in this range. For a feature to be selected by either method, the addition or removal of the feature had to result in a correspondence score of under 90% — that is, the change in the feature set needs to induce more than 10% of the passenger labels to switch to be considered significant. For each k-value, we also randomized the order of the inputted feature list at least five times, to ensure our results were not biased by the order of features tested in these sequential methods.

Bus share and peak share were universally selected, followed by active range (selected by 88% of runs), weekend share (81%), transfer share (69%), and active days

(67%). Average weekly journey intensity was only selected in 33% of runs, so we excluded it. On runs where active weeks was tested before active days, active weeks was often chosen instead, suggesting that we could have used either active days or active weeks as a final feature. However, we chose active days because the correlation between active range (a feature we will definitively keep) and active days had a lower correlation of $r = 0.75$, compared to $r = 0.93$ between active range and active weeks. As discussed in Section 3.4.1, it is preferable to run k-means with less correlated features. Thus, our final six selected features were bus share, peak share, active range, weekend share, transfer share, and active days.

4.2.2 Selected k-value

Given the selected features, we ran k values from 3 to 10 in order to assess the optimal k with the elbow method and DBI. In the elbow method, we looked for the k beyond which the drop in inertia per incremental cluster slowed. This occurred around $k = 5$ in Figure 4-5a. In DBI, we look for the k that minimized the DBI score, which in this case occurred at $k = 7$ (Figure 4-5b) indicating that the features we chose were successful at creating 7 clearly distinguishable clusters. We therefore proceeded with seven clusters for multi-day riders.



(a) Elbow plot

(b) Davies-Bouldin Index

Figure 4-5: Elbow method, Davies-Bouldin Index results for k selection among multi-day riders

4.2.3 Multi-day rider clusters

Given the above hyperparameters, we ran k-means to generate seven distinct clusters for multi-day riders. We present these below arranged by descending passenger count. We describe these clusters in terms of the "centroid" value, which is the central or representative point against which each datapoint is clustered. The heatmap table shown previously in Figure 4-1 summarizes these feature centroids and, to help contextualize the relative operational importance of each cluster, it also indicates the size of each cluster in terms of riders and journeys taken. Figure 4-6 visualizes each cluster by its centroids, in order to provide a snapshot of what makes each cluster unique. Values for each feature are normalized to [0%, 100%] so that we can plot all features along the same set of radar chart axes (for variables that were not already percentages, normalization means that a point taking on, for example, the minimum active days value would map to 0%, and a point taking on the maximum value would map to 100%).

Finally, to provide a sense of the compactness and distinctness of each cluster, we visualize the variance of each feature within each cluster using violin charts in Figure 4-7. Wider areas of each violin is where more cluster members are concentrated; the thicker, dark bar in the middle of each chart represents the interquartile range (IQR, i.e., the box portion of a box plot).

Cluster #1: Frequent bus riders

The frequent bus riders cluster is the largest, capturing 11% of baseline period passengers, equivalent to 172,552 people. Given that these passengers were also frequent users, the cluster contributed by far the largest share of journeys at 35%, equivalent to 6.7 million journeys. Together, these statistics suggest the out-sized importance of this bus-oriented group in terms of ridership, system operations, and potentially also system revenue.

Riders in this cluster were heavy bus users. At the centroid, 75% of trip stages were by bus. Figure 4-7 indicates that there was a wide spread in bus share within

Figure 4-6: Characterization of multi-day rider clusters by feature centroids (normalized to 0-100%)

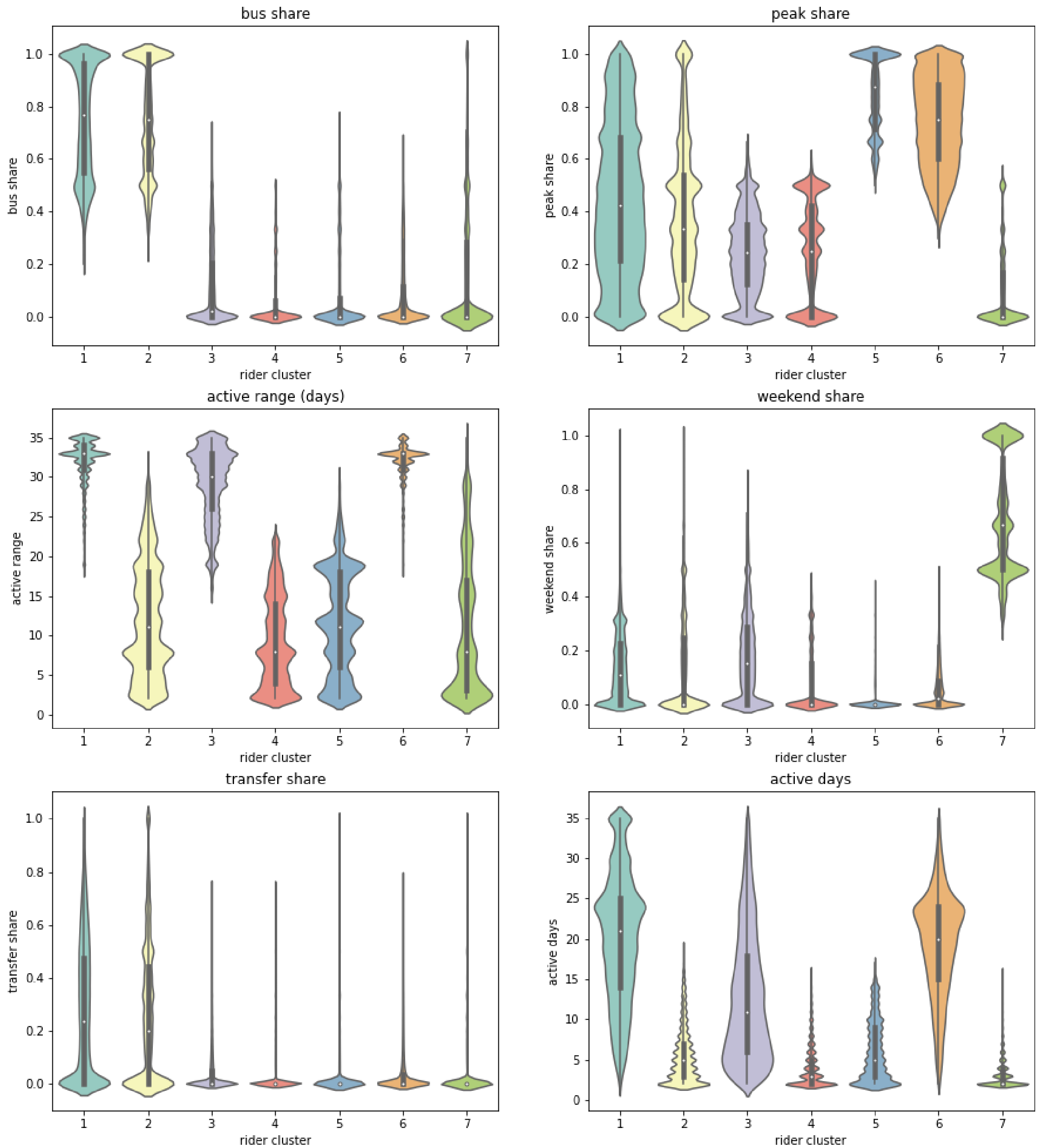


Figure 4-7: Variation of features within and between multi-day rider clusters

this cluster (18% to 100%), but the majority of the distribution was above 50%, far outstripping the bus usage of clusters 3 through 7 where the majority of passengers took under 10% of their trip stages by this mode. Despite the high bus usage, bus-related transfers were only moderately common (27% of journeys at the centroid and with the distribution weighted towards the lower end of the range).

The time of day during which passengers in this cluster entered the MBTA system is not dominated by rush hour commute trips. Instead, this cluster displays some of the greatest diversity in the time of day of transit use. The centroid value for the share of journeys begun during peak hours is moderate at 45%, indicating that the average frequent bus rider had some peak-period commutes. However, Figure 4-7 indicates that passengers in this cluster ranged fairly evenly between having no rush hour trips to only riding during rush hour. The only other cluster with somewhat comparable timing diversity was cluster 2, the occasional bus riders. The rail clusters (#3-7), on the other hand, show stark time of day specialization among rail riders, with frequent and occasional rail commuters (clusters 6 and 5 respectively) heavily skewed towards peak trips, while the remaining rail-oriented riders tended to avoid those busy commuting hours. The moderate amount of peak travel in this frequent bus riders cluster, alongside the fact that the centroid value for the share of weekend journeys is very low at 14%, suggests that this cluster undertakes substantial off-peak weekday travel and displays wide temporal range in transit usage.

This cluster also hosts some of the most intensive MBTA riders. The centroid value of active range and active days were 31.98 days and 20.12 days respectively, indicating consistent usage throughout nearly the full baseline period. Active range also had relatively little variation across cluster members, according to Figure 4-7. The number of active days within that window varied widely, but is more heavily weighted towards the higher end of the distribution. The only other cluster with a comparable usage intensity as measured by both active range and active days is cluster 6, the frequent rail commuters.

Cluster #2: Occasional bus riders

This cluster captured 10% of baseline period passengers, equivalent to 152,429 people. Because its riders took transit more occasionally, this relatively large passenger pool only contributed 8% of baseline journeys, or 1.6 million.

Similar to the frequent bus riders, this cluster is heavily bus-oriented. At the centroid, 76% of journeys were taken by bus. Figure 4-7 suggests that the bus share is dominated by higher values but is still quite varied among passengers of this group, as it was for the frequent bus riders. The centroid value for transfer share (26%) as well as its distribution were also similar to frequent bus riders.

These two bus-dominated groups were additionally similar in displaying wide variation in the share of peak trips, though among occasional bus riders there were clearer modes around 0%, 50%, and 100% peak travel. At the centroid, the share of peak journeys is 37% which is even lower than for frequent bus riders. Occasional bus riders also tend to engage with the system on weekdays rather than weekends, with only 14% of journeys taken on weekends at the centroid. Thus, this cluster again suggests a wider dispersion to the travel schedules of bus passengers, with plenty of off-peak weekday travel.

The key differentiator between this bus-oriented cluster and the previous one is transit ridership frequency. Passengers in this cluster had fewer active days overall (5.48 days at the centroid), and were also active over a smaller window (the centroid value for active range was 11.91 days).

Cluster #3: Fairly frequent weekday off-peak rail riders

This is the largest group of rail-oriented passengers, capturing 9% of baseline period riders, equivalent to 134,790 people. Because of the fairly frequent transit usage of this cluster's members, it is responsible for an even larger share of all journeys taken during the baseline period—17% or 3.2 million journeys. The size of this cluster attests to its importance for MBTA operational planning.

Fairly frequent weekday off-peak rail riders were more heavily tilted toward rail use than the bus-oriented groups were toward bus. Bus only contributed 11% of all trip stages at this cluster's centroid. More broadly, Figure 4-7 suggests the rail-oriented clusters (#3-7) across the board had much lower cross-engagement with buses than the bus-oriented clusters had with rail.

Temporally, travel in this cluster is neither peak-oriented nor weekend-oriented, with the centroid value for share of peak journeys at only 24%, and for weekend share at only 18%. This suggests that most travel in this cluster is instead off-peak weekday usage. Figure 4-7 also indicates that there is much less spread in peak share here than what we've observed so far with the two bus clusters—the entire IQR of peak share falls within 12% to 37%.

The passengers in this cluster are active over almost the entire baseline period, with a centroid value of 29.25 days for active range. The number of days they actually ride transit within that window, 12.87 days at the centroid, is notably lower and only 45% of the centroid active range value. This is still substantial usage frequency though, reflecting one trip nearly every other day on average over most of the baseline period.

Cluster #4: Occasional weekday off-peak rail riders

This cluster is one of three that describe the dispersed travel patterns of occasional rail users. It captures 8% of all baseline period passengers, equivalent to 127,368 passengers. Because its members rarely engage with the transit system, its contribution to the total baseline journey count is even lower at 5%, or only 918,451 journeys over the five-week window.

This cluster is extremely skewed towards rail, with only 6% of trip stages taken by bus at the centroid. It focuses on weekday off-peak users, with only 25% peak journeys and 7% weekend journeys at the centroid. The usage rates are low, with the centroid exhibiting a short 9.18 days of active range and only 4.07 in actual active days.

Cluster #5: Occasional rail commuters

The larger of our two clusters focused on rush hour travel, the occasional rail commuters segment captures 7% of baseline period passengers, equivalent to 110,398 riders. It also represents a similar proportion of total journeys, at 6% or 1.2 million.

At the centroid, 85% of journeys in this cluster began during peak hours. Looking at the broader distribution pictured in Figure 4-7, there is a clear mode at 90% or more rush hour travel. There is nearly no bus use, and both the active range and active days are low (11.52 and 6.14 days at the centroid respectively).

Cluster #6: Frequent rail commuters

Though this cluster only represents 7% of all riders or 109,225 people, its intensive transit usage means that it captures an out-sized share of total journeys—21% or 4.0 million journeys. The high volume of this cluster during the most congested hours of the day attests to its importance for MBTA operational planning—as does its relationship to corporate sponsored fare products, which we will discuss in section 4-18 below.

This cluster, like the occasional rail commuters, is heavily skewed away from bus and towards peak-period travel. At the centroid, only 8% of trip stages are taken by bus, with 74% of journeys occurring during peak hours. The weekend share of journeys is only 5%. The intensity of usage is as high as for the frequent bus riders group, with the centroid active range at 31.85 days and active days at 19.36.

Cluster #7 Occasional weekend rail riders

This is the smallest multi-day rider cluster in terms of both rider count and journeys, contributing only 5% of riders (73,976 passengers) and 2% of journeys (391,277 journeys). It is predominantly rail, with only 15% of trip stages taken by bus at the centroid. It is the only weekend-oriented multi-day cluster, with 68% of all journeys taking place on the weekend at the centroid compared to under 18% for all other multi-day rider clusters. There is little peak-period travel in this cluster—the centroid value is 10% and the overall distribution has very little mass at upper values (Figure 4-7). The number of active days is the smallest of any of these multi-day clusters, with a centroid value of 3.2 days spread over an active range of only 11.08 days. This segment likely reflects those who use transit mainly for leisure weekend activities.

In summary, bus clusters are clearly distinct from rail clusters along the modal dimension largely because rail-oriented riders have low engagement with buses. Within each mode, clusters are differentiated by usage frequency. The baseline clustering results indicate greater dispersion in passenger habits for rail use than for bus use, given that bus riders are grouped into two large clusters while rail users are spread out over five smaller ones. Bus-oriented multi-day riders are a smaller subset of the data than rail-oriented ones are to begin with (324,981 versus 555,757 riders), but there are still proportionately fewer bus clusters than rail ones even after adjusting for passenger numbers. For rail-oriented riders, there are distinct sub-groups that focused on certain times of the day or week, which k-means was able to distinguish and return as separate clusters. Occasional multi-day rail users form the largest number of distinct behavioral groups, with some focused on commuting during rush hour, others on weekday off-peak travel, and still others on weekend activities.

4.3 Baseline transit rider clusters: Single-day riders

4.3.1 Selected clustering features

For single-day riders, several features previously selected for multi-day users can no longer be included due to lack of variability. These include the active range and active days features, each of which are now one day by definition. Further, there is little variation in mean journeys per active week. In general, raw features for each card show much more extreme values for single-day riders because of the lack of averaging over multiple days; for example, weekend trips is actually categorical, since a single-day user is either riding on the weekend or not (Figure 4-8). Only peak share and bus

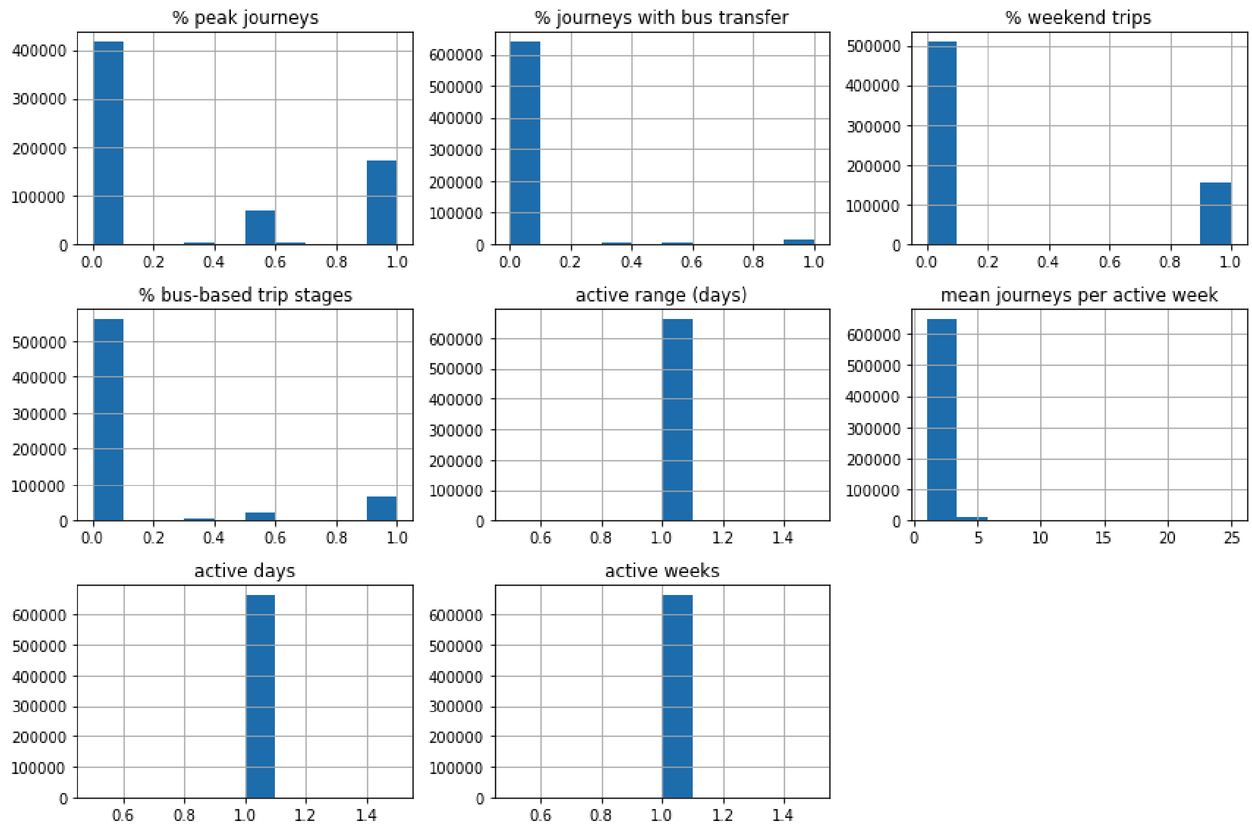


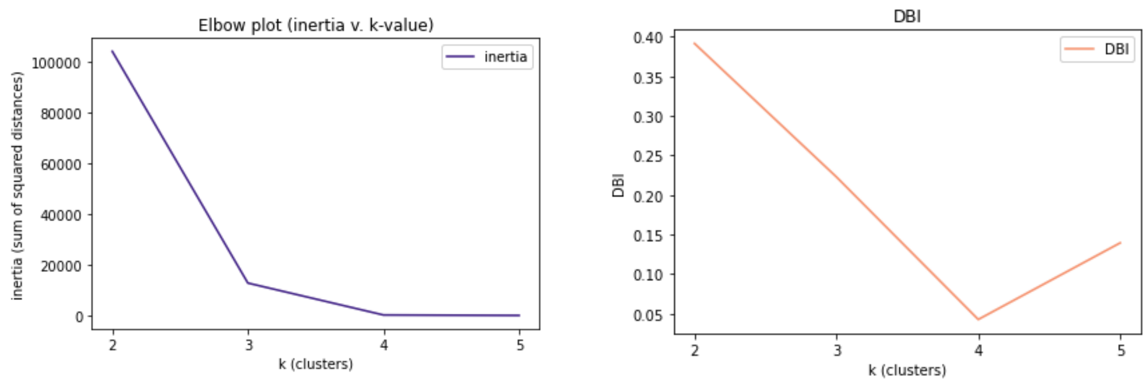
Figure 4-8: Distribution of candidate clustering features for single-day riders

share are more split between a small number of values. This already suggests that only a few features would likely to be useful for partitioning single-day passengers.

The feature selection process was conducted with greedy and backward selection, for k values from two to five as suggested by a preliminary elbow plot. Peak share and weekend share were selected in 85% and 92% of runs, while all other tested features were selected at most 46% of the time. These features were only moderately correlated (r = -0.4). We thus proceeded to clustering using only these two features.

4.3.2 Selected k-value

The elbow plot and Davies-Bouldin Index suggested between three and four clusters as illustrated in Figure 4-9. We produced initial results with both three and four clusters to assess the cleanliness and interpretability of the results. Either k value produced clean partitions, but we ultimately proceeded with three clusters because



(a) Elbow plot (b) Davies-Bouldin index

Figure 4-9: MBTA tap-in data by mode and by source

the fourth cluster was extremely small (under 10% of single-day riders), within a subset of the data that was otherwise cleanly partitioned by a small number of very large clusters. The fourth potential cluster covered a limited group of weekday riders with more mixed peak/off-peak schedules.

4.3.3 Single-day rider clusters

Running k-means with the above hyperparameter decisions produced three clusters that were divided solely by the timing of transit usage. These clusters' members are rarely trackable over time due to their transient engagement with the system, and as shown previously in Figure 4-1 they also are the smallest contributors to the volume of journeys taken on the MBTA system (i.e., they have little impact on MBTA operations and revenues). Thus we will not concentrate on profiling these riders in detail. However, it is still essential to create and introduce these clusters upfront, because during COVID-19 a notable volume of passengers reduced their transit usage and took on the behavioral patterns observed in single-day rider clusters.

Figure 4-10 shows the distribution of peak and weekend journey shares among members of single-day clusters, alongside the distributions for multi-day clusters for comparison.

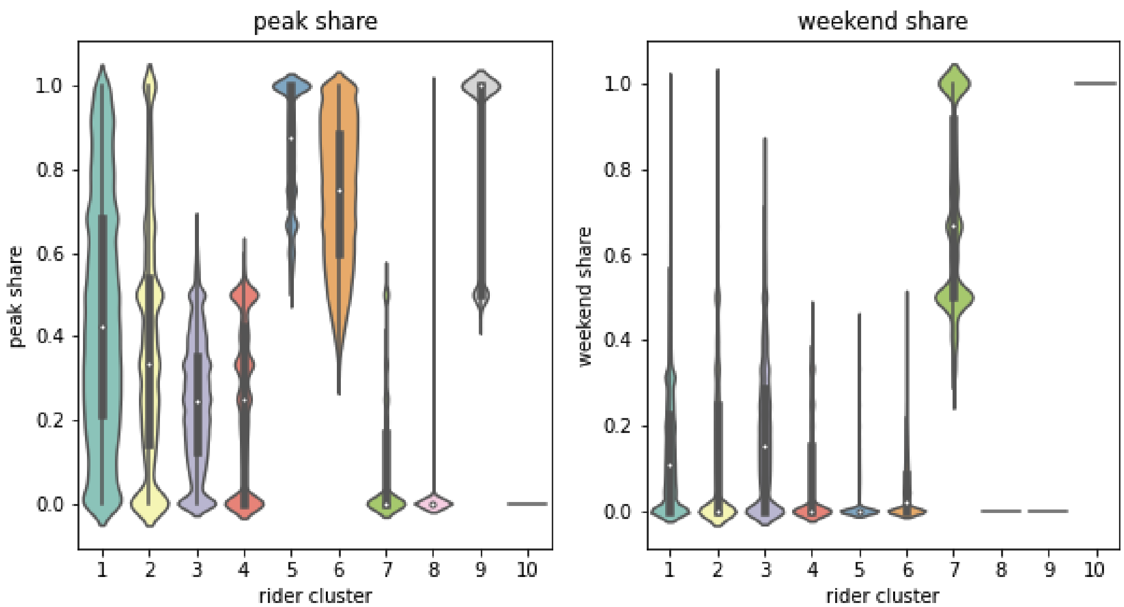


Figure 4-10: Variation of single-day clustering features within and between rider clusters

Cluster #8: Single-day off-peak weekday riders

This cluster was by far the largest among both multi-day and single-day riders, capturing 18% of all baseline period cards or 271,676 passengers compared to only 11% for the largest multi-day cluster of frequent bus riders. However, its operational and revenue significance for the T is small as it contributes only 2% of journeys—compared to 35% for frequent bus riders. At the centroid, 0% of journeys were on the weekend and there were nearly no peak journeys (0.66%). Thus, cluster members only took trips during off-peak hours of one weekday. There is also no variation in the weekend share feature in this cluster; along the peak share feature, there is a long tail of higher values. Lastly, as Figure 4-1 indicates, this group like all single-day rider groups is dominated by rail use, with the average bus share at only 14%.

Cluster #9: Single-day peak riders

Contributing 15% of all riders (236,792 people) but only 2% of journeys, this cluster represents passengers who took transit during rush hour for a single day. The centroid value for peak journey share was 85%, similar to the shares for the multi-day rail commuter clusters, and weekend share was at 0%.

Cluster #10: Single-day weekend riders

Like the other single-day clusters, this was one of the largest clusters by rider count, but the absolute smallest by journey count among all multi-day and single-day clusters. It covered 10% of all cards or 155,101 passengers—similar in passenger volume to the frequent and occasional bus rider clusters—but contributed only 1% of trips. It contains single-day riders that utilized the system on a weekend day, with a weekend share of 100% at the centroid and a corresponding peak share of 0%.

In conclusion, single-day riders mostly took rail and were small contributors to demand for transit services. They were partitioned into clusters fairly cleanly by whether they were weekend, peak, or weekday off-peak users.

4.4 Baseline model validation: Temporal robustness checks

As illustrated in our methodology flowchart in Figure 3-1 of the previous chapter, our next step after clustering was to validate the baseline k-means model by checking its temporal robustness. To do so, we assess the stability of the established clusters between seasons and between years by 1) training k-means models from scratch on winter, spring, summer, and fall 2019 data, then 2) using correspondence scores to compare the clustering results of these new models against the results we receive when we apply the baseline model to these other periods of data (see Section 3.5). The main focus here is whether the structural composition of the rider clusters is still similar to what we found in our baseline winter 2020 period.

Figure 4-11 shows the aggregate trip stage count by mode from ODX across the months of 2019 considered for our temporal robustness check. Usage appears to trend up slightly from winter to spring, then drop over the summer holidays before jumping back up in September for the fall. This uptick, which occurs for both bus

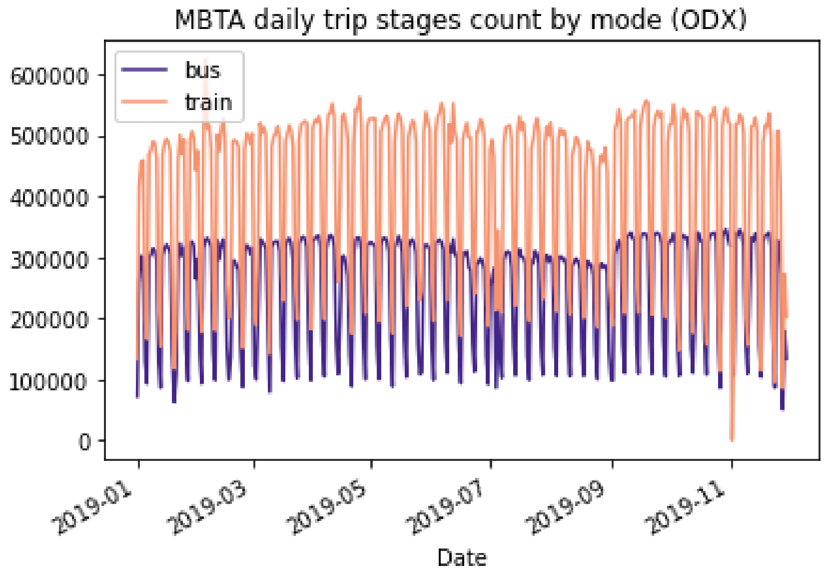


Figure 4-11: Rail and bus ridership in 2019, from trip stages recorded in ODX

and rail, may be due to the usual return to work and school, given metropolitan Boston's student-oriented nature. Figures 4-12 shows in greater detail the rail and bus ridership patterns for each of the four specific time windows sampled. Weekend travel by train is notably lower in the winter and colder parts of fall than in the other seasons; the weekend gap in ridership between bus- and train-based modes is also closer together during these periods.

Overall, the features selected for multi-day and single-day clustering and the number of finalized clusters appear consistent over the time periods examined. The centroid tendencies (i.e., the characterization) of the clusters are also *qualitatively* similar across all time periods tested. For example, for each time period there are always two clusters with much higher bus shares than all other clusters, two multi-day rider clusters with extremely high peak journey shares compared to all others, etc. As a result, the correspondence scores of each cluster is also high—above 90%—across all clusters and all 2019 seasons tested.

However, the actual centroid *values* for multi-day clusters differed noticeably across seasons in a few clusters. Figure 4-13 plots the *difference* in normalized centroid value of each multi-day cluster between the baseline 2020 period and each season of 2019. Thus, for example, the "Cluster #1" radar chart within that figure displays the

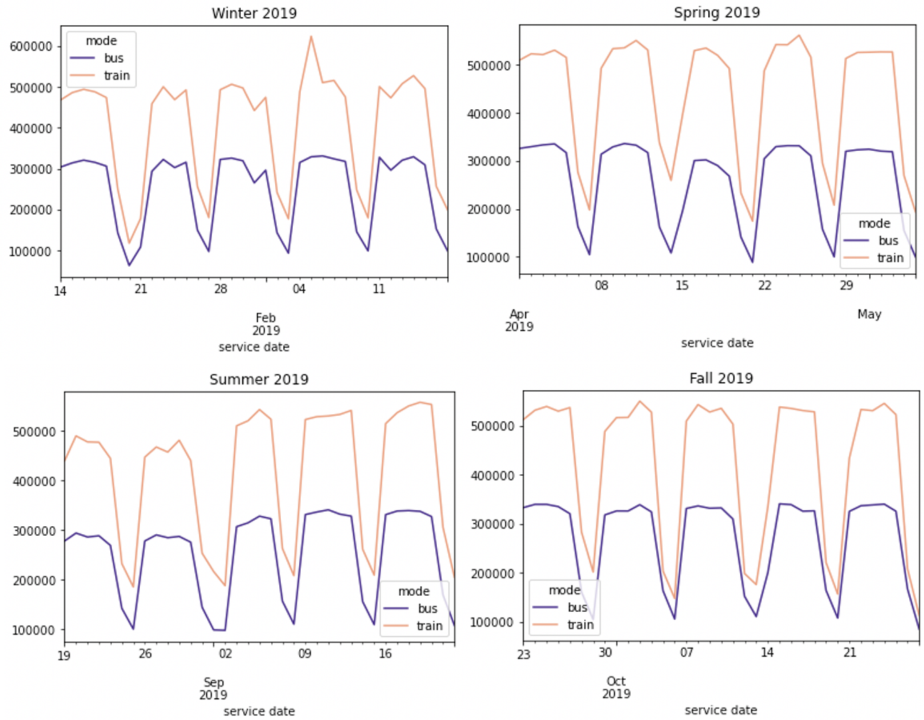


Figure 4-12: Trip stages by mode in 2019 time periods used for temporal robustness checks (ODX)

baseline 2020 centroid values minus the winter 2019 centroid values for each feature; it also does the same for spring, summer, and fall 2019. The winter 2019 centroid differences are depicted in solid color since it is the season expected to match the closest to baseline 2020 values. The spring 2019 centroid differences are dotted lines; summer 2019 dashed lines, and fall 2019 the dot-dash. Single-day cluster centroids were not included as they were all nearly identical between the robustness check period models and the baseline 2020 model. Figure 4-13 shows that in general, winter 2019 and baseline period winter 2020 compare fairly well, with no centroid showing extreme differences. The fairly frequent weekday off-peak rail riders, occasional weekday off-peak rail riders, frequent rail commuters, and occasional weekend riders are also similar, with the difference between baseline and 2019 seasonal model centroid values differing by at most 4 ppt. On the other hand, frequent bus riders, occasional bus riders, and occasional rail commuters showed notable centroid value differences for spring and sometimes also one other season.

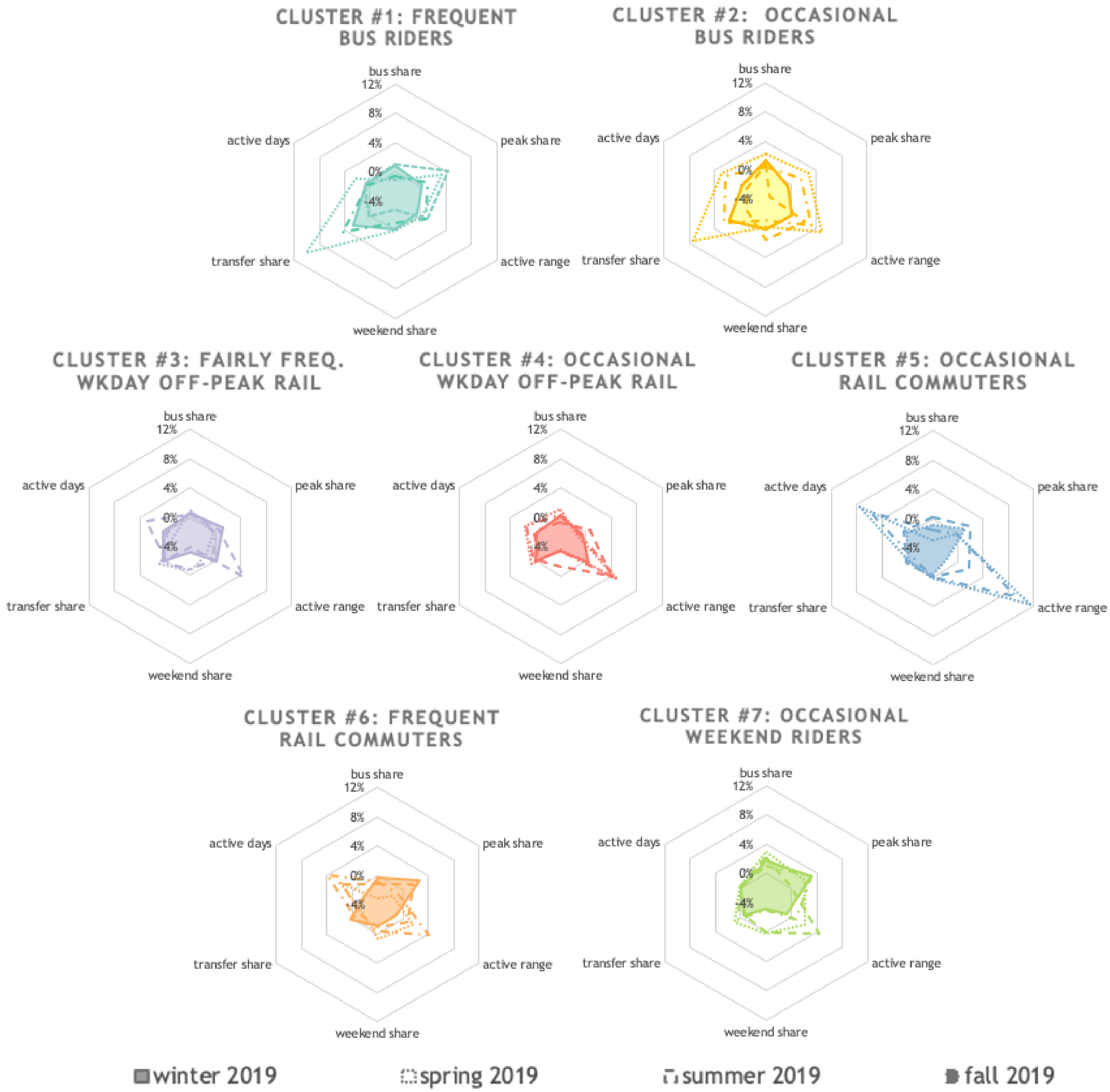


Figure 4-13: Centroid differences between 2020 baseline model features and 2019 seasonal model features

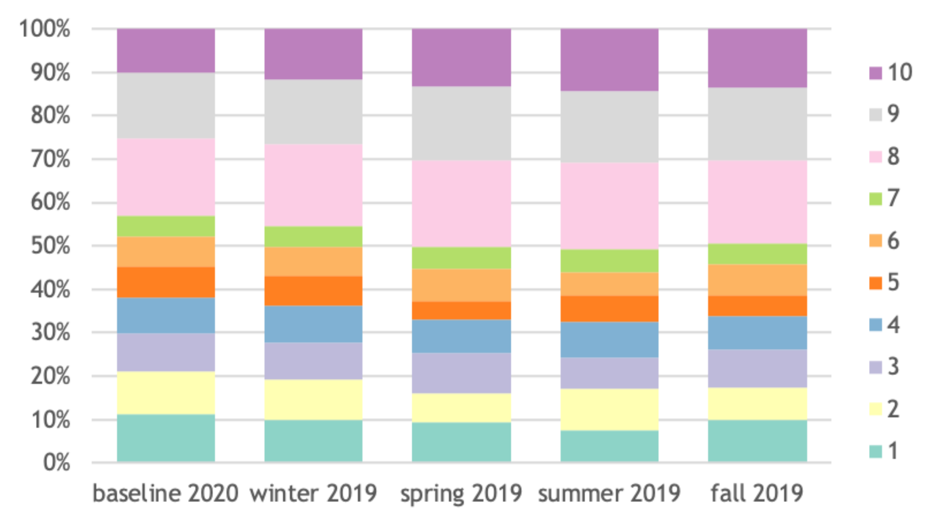


Figure 4-14: Cluster composition for baseline and winter, spring, summer, fall 2019 passengers

Finally, Figure 4-14 shows that there are seasonal differences in the proportion of passengers falling into each behavioral cluster. While baseline 2020 and winter 2019 saw extremely similar cluster distributions, the warmer seasons saw greater shares of all three single-day clusters. Cluster #5, occasional rail commuters, was also a smaller share of the passenger pool in the shoulder seasons.

In summary, the temporal robustness check finds that the baseline winter 2020 k-means model applies very well to the winter ridership of an adjacent year. It also performs fairly well on other seasons, but the distribution of passengers across clusters and the exact centroid value denoting representative cluster behavior sees greater variation across seasons. Detailed results for each of the four 2019 seasons tested are discussed in the sub-sections below.

4.4.1 Winter 2019: January 14 - February 17

For winter 2019, the data contained 22.9 million trip stages by 1.65 million riders, similar to what was observed during the baseline winter 2020 period. The six clustering features selected by the greedy and backward selection methods when modelling multi-day riders during the baseline were also selected during the winter 2019 modelling. In addition, backward selection (but not the greedy method) recommended the use of average journeys per week for the winter 2019 data. However, we leave this candidate out from the final feature set because it is highly correlated ($r = 0.72$) to active day, the latter of which was selected by both methods.

The Davies-Bouldin index again recommends that $k = 7$ clusters for multi-day users. After running k-means for this hyperparameter value, we evaluate the similarity of this clustering solution to the baseline period's and find that correspondence scores are above 95% for each cluster except the occasional weekday off-peak rail riders, which scored 93%. The worst correspondence score belonged to the smallest cluster, which is to be expected since estimates of population behavior made based on smaller samples show greater variance. If we run an overall correspondence score at the full sample level instead of at the cluster level—i.e., we calculate one correspondence score over all sampled passengers, instead of seven scores one for each cluster—we find that the score is satisfactorily high at 98%.

Qualitatively, the centroid values of the stability check model also align with that of the baseline model, with no difference greater than 4 ppt (Figure 4-13). This suggests that the detected clusters represent similar behavioral patterns to those observed in the baseline period. Further, the distribution of passengers across clusters—i.e., the relative cluster sizes—are different from baseline figures by at most 2 ppt.

For single-day passengers, weekend and peak share are again the selected features, and k is again set at three clusters. Comparing classifications between the model trained on the stability check dataset, versus the model trained on the baseline dataset, we find that correspondence scores rounded to 100% for all single-day clusters, testifying to the clean partitions that can be made among single-day riders based on the timing of their journeys. The share of riders falling into each cluster is similar to baseline 2020 values as well.

4.4.2 Spring 2019: April 1 - May 5

Here, 24.6 million trip stages were recorded by 1.88 million riders, reflecting a moderate increase in trip stages (9%) alongside a large expansion in riders 22% compared to baseline 2020 levels. CharlieCards are slightly less prevalent than during the baseline winter period, coming in at 49% of trip stages instead of the the 52% as in winter 2020. This may be due to the uptick in tourism and leisure trips among those who generally do not depend on transit, as the weather warms. This qualitative difference appears in the pattern of correspondence scores we find across clusters and the distribution of spring-time passengers across these clusters.

The feature selection process is similar to winter 2019, with average journeys per week again recommended by the backward selection method in addition to the baseline 2020 model features. Since we are clustering at the passenger level and passenger count is 22% higher than in the baseline period, there is increasing justification to consider adding another feature as the dataset becomes larger and the chance that additional distinct clusters exist increases. We do not do so at this time, but recognize that there may be a need to do so if both the greedy and backward selection methods consistently pick this candidate feature for other high-volume time periods.

We again select seven multi-day and three single-day clusters and find that the correspondence scores are above 90% for every cluster. The sample-wide individual-level correspondence score ignoring cluster distinctions is 97%, only 1 ppt below the score found for winter 2019. These results seem to indicate inter-seasonal model stability. The two clusters with the lowest correspondence scores are frequent rail commuters (90.35%) and frequent bus riders (93.88%), both of which likely include significant work-based travel.

Looking at the differences in centroids between spring 2019 and baseline 2020 model results (Figure 4-13), the frequent rail commuter cluster characteristics appear highly similar between models trained across the two time periods. The frequent bus riders cluster centroid, however, shows notable differences from baseline to the spring. The baseline period's transfer share is 11 ppt higher and its peak share 4 ppt higher. Occasional bus riders and occasional rail commuters also show notable centroid differences between spring 2019 and baseline 2020 despite having higher correspondence scores (94% and 97% respectively). For the occasional bus rider cluster, the baseline model's transfer share is 8 ppt higher than the spring 2019 model's, and active range is 5 ppt higher. For occasional rail commuters, the baseline model's active range is 11 ppt higher than the spring model's and active days 7 ppt higher. Both these deviations from baseline winter 2020 centroids (and winter 2019 centroids) may reflect an uptick in alternative modes like walking and biking as the weather warms.

As Fig 4-14 indicates, single-day riders also contribute 7 ppt more to the overall ridership pool in the spring than in the winter 2019 or 2020 periods, potentially reflecting more ad-hoc ridership by locals and more visitors to the Boston Metro area with the end of winter. The clusters that are more likely to reflect work-based riders or transit-dependent riders fell as a share of the overall pool–occasional rail commuters fell from 7% to 4%, and both the frequent and occasional bus rider groups fell 2-3 ppt each. Overall though, despite some change in the cluster composition and in a handful of centroid values, the correspondence scores suggest a relatively robust application of our baseline k-means model to spring 2019 data.

4.4.3 Summer 2019: August 19 - September 23

The summer 2019 dataset was substantially larger than the baseline winter 2020 dataset, with 2.02 million riders (31% above baseline winter 2020). However, the number of trip stages is only 3% higher at 23.4 million, already suggesting that much of the new cards in the system may be single-day or occasional users, perhaps tourists and other leisure riders who do not consistently return to the system. This is further suggested by the smart card meta-data, which indicates that a slightly smaller share were CharlieCard users than is the case in the baseline (47% rather than 52%). The distribution of active range among the multi-day riders also shows lower density at the higher end (i.e., fewer people are riding over 30 active days or more).

The features and k-value selected match the results from the baseline winter 2020 model, for both the multi-day and the single-day rider pools. Results again show that 8% more of passengers fell into the three single-day rider clusters than had been the case in baseline winter 2020, and 5% more than in winter 2019. Meanwhile, the frequent bus cluster was down from 11% to 7.4% of the total—a 50% drop. The fairly frequent weekday off-peak rail cluster is down 20% from 9% of all riders to 7%, while the frequent rail riders cluster is down from 7% to 5%. In addition to workers and students going on vacation in the summer, the better weather can also support more biking and walking, which can help explain the drop in ridership among the more frequent rail riders.

Correspondence scores indicate stability between the application of the k-means model trained on winter 2020 baseline data and the model trained on the summer 2019 data itself. Overall ignoring cluster boundaries, 98% of all passengers are assigned the same cluster by both models. At a cluster level, correspondence scores are above 94% for all but the frequent rail commuters (91%) and occasional rail commuters (90%). Journey counts suggest these groups show reduced activity over the summer than over the winter. They therefore may have exhibited behavior that shifted them into one of the five less-frequent rail user groups. Since this winter-summer comparison will provide the "background" or "natural" behavioral shift against which we compare the pre-COVID baseline winter 2020 versus COVID-period summer 2020 ridership patterns, we will need to keep in mind this greater cluster shift among rail commuters evident in the 2019 data.

The centroids from the summer 2019 model are similar to that from the baseline model, with no differences greater than 4 ppt in absolute value. Across all clusters, the baseline centroid value of active range is 1 ppt to 4 ppt higher than summer 2019 centroid values; this means there is a slight shift in the summer distribution towards the lower end of active ranges seen in the sample, which is reasonable in warmer weather months with more visitors using transit and fewer regular riders. The occasional and frequent rail commuter clusters' peak ridership is also 2 ppt higher in the baseline than the summer model. Overall, the centroid values and correspondence scores suggest that our k-means model trained on the baseline 2020 data is applicable towards summer ridership, though we need to be aware that there is a natural shift in cluster composition towards single-day ridership.

4.4.4 Fall 2019: September 23 - October 27

Ridership in fall 2019 was substantially larger than during the winter, as both students and workers returned from summer vacations to full-time work or study. The robustness check dataset for this period includes 24.3 million trip stages (a significant 7% above baseline 2020) taken by 1.9 million riders (23% higher than the baseline). The active range distribution resembles the winter baseline period's in that the mode at 30+ days has become more prominent again after the summer.

Feature selection performed upon the fall 2019 robustness check period again produced similar results as during the baseline, except the greedy method suggested the share of journeys with bus-based transfers may not be necessary. This is a detail to follow up on for future model development, since in our original baseline clustering results transfer share patterns mainly reflected patterns in the share of bus transfers, perhaps because 1) our data fails to capture rail-rail transfers and 2) rail-based riders have very low interaction with the bus system, limiting their opportunity for bus-rail transfers. Separately, while larger spring volumes led our feature selection process to suggest an additional candidate feature to better partition clusters, this was not the case for fall 2019 despite the notably higher passenger volumes.

For multi-day riders, k is again set at seven based on the Davies-Bouldin Index and elbow method, while k is set at three for single-day users. The overall correspondence score ignoring cluster boundaries again indicates that 98% of passengers are classified into the same clusters by both the baseline and the fall 2019 models. At the cluster level, most clusters have correspondence scores of over 94%; the exceptions are the frequent rail commuters (91%) and occasional bus riders (93%).

The fall 2019 model's centroids deviate most clearly from the baseline 2020 model for occasional rail commuters. The baseline model's active range is 9 ppt and active days is 6 ppt higher than for the fall 2019 model (Figure 4-13). Among frequent rail commuters, these deviations are 4 ppt for both. For occasional bus riders, the baseline centroid values is not much greater than the fall model's along any feature, but all features show noticeable deviation with active range, active days, peak share, and transfer share being the most prominent. Otherwise, the centroid values were fairly in line with the baseline model's.

One potential factor in the lower active day and active range seen among commuters could be weather–early fall weather in Metro Boston is much more amenable to biking and walking than in the winter (or even most of spring), so riders with shorter commutes or commutes along bike- or pedestrian-friendly routes may choose this option. This would directly reduce active days. To impact active range, this would have to occur systematically at the start or end of the dataset time period, but this is not unreasonable given the cooler weather rolling in as fall deepens.

The overall cluster composition again favors single-day clusters compared to the baseline by 6 ppt. This comes at the expense of multi-day user clusters, especially the occasional bus riders and occasional rail commuter clusters which each experience 2 ppt decreases in their share of the total rider pool.

All together, the results suggest that the k-means model trained on the baseline 2020 dataset is fairly stable across seasons, and that it may be stable across neighboring years. The differences noted in the winter versus summer cluster composition again emphasizes the need to compare our churn analysis in 2020 against the "background" winter-to-summer churn during a normal year, in order to separate the COVID-19 effects from natural seasonal churn.

4.5 Cluster interpretation with optimal decision trees

As discussed in Chapter 3, one objective of this thesis is to interpret our baseline clustering and subsequent COVID-19 era churn results in ways that facilitate policy discussion. While most of the policy work is conducted in Chapter 6, this section covers cluster interpretation with optimal decision trees (OCT) because methodologically, it is a direct extension of our quantitative k-means clustering work. Further, we conduct this baseline interpretation step prior to considering COVID-19 churn so as to improve our understanding of key factors driving transit behavior during typical times and bring that mechanistic understanding to the upcoming churn analysis.

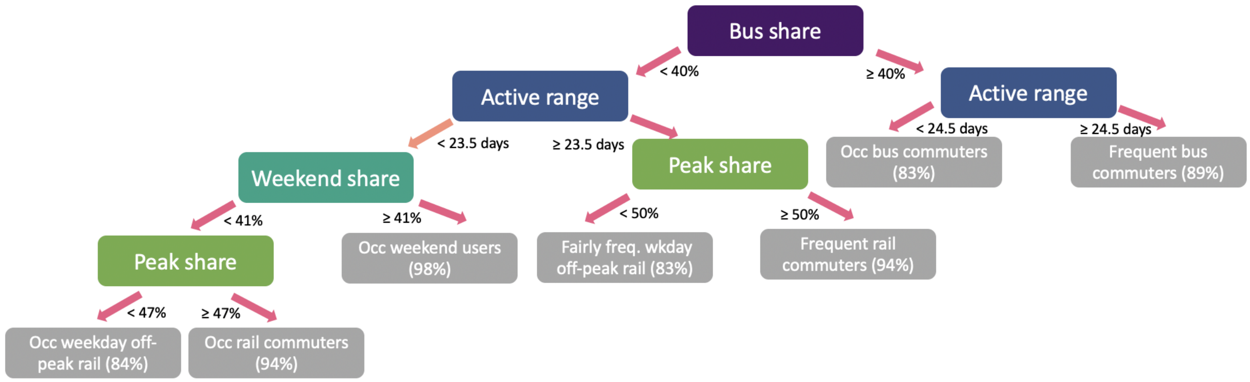


Figure 4-15: OCT for multi-day riders using clustering features

Optimal decision trees (OCT) were applied separately to multi-day and single-day passengers, following the division in our clustering results. Overall, the results confirm the importance of modality, followed by usage intensity and time of day of transit usage, in distinguishing between passenger groups.

4.5.1 OCT for multi-day riders

We first ran OCT using the multi-day rider dataset, applying the cluster values from k-means as labels to convert this unsupervised dataset into a supervised one. Grid search was used to tune the complexity hyperparameter α and the tree depth for this univariate OCT. We attempted two grids: complexity [0.1, 0.01, 0.001, 0.0001] with depth [1, 5] and complexity [0.1, 0.05, 0.01] with depth [1, 10]. Balancing tree legibility and interpretability against the classification error rate, the finalized tree for multi-day riders was found at $\alpha = 0.01$ and tree depth = 4. This tree is shown in Figure 4-15, with the branching nodes in color and leaves with cluster labels in grey.

The resulting classification accuracy is fairly high, with at least 80% of points classified correctly in each leaf. The first split is made on modal preference, that is, the share of trip stages a passenger conducted by bus.[1] Subsequently, active range, which is a measure of usage intensity, tops the list among both bus-heavy and train-heavy riders. Among bus-oriented riders, this additional feature is enough to deliver

[1]The cut-off OCT selected for splitting bus-oriented from rail-oriented riders was 40%, which is near the bottom of the range for the two bus clusters (Figure 4-7). Yet because the bus share was so low among nearly all rail-oriented riders (<15%), the 40% cutoff allowed a fairly clean split.

the two bus clusters detected by k-means, though the accuracy is more modest at 83% for occasional bus riders.

There are more than twice as many rail-oriented rider clusters, so a larger number of features are needed to distinguish between them. Among rail riders with high active range ($\geq$ 23.5 days out of 35 days in the baseline period) the share of journeys taken during peak weekday hours is the main distinguishing feature that allows us to clearly separate out frequent rail commuters (94% accuracy) from fairly frequent weekday off-peak rail riders (83% accuracy).

On the other hand, among rail riders with limited active range ($<$ 23.5 days), the share of journeys conducted on weekends is the next branching point. Those with heavy weekend travel ($\geq$ 41% of all journeys) were labeled occasional weekend riders. Those preferring weekday transit travel are subdivided by the share of peak period trips into occasional weekday off-peak rail riders versus commuters. The order of features used in this OCT appears to confirm the importance of modality, usage intensity, and timing of transit usage in partitioning multi-day riders.

The share of journeys with transfers and the number of active days were not needed for partitioning to the level of accuracy achieved in this OCT. The utility of transfer share as a clustering feature does appear limited from our previous analysis as well, due to both the lack of data for capturing rail-rail transfers and due to the the low percentage of journeys with bus-based stages taken by rail-oriented passengers.

Though the accuracy of the OCT in categorization is strong as a whole, the worst performance were seen in the three categories that include substantial off-peak weekday travel. These are the occasional weekday off-peak rail, fairly frequent weekday off-peak rail, and occasional bus riders who were shown in Figure 4-6 and 4-7 to have a wide mixture of peak- and non-peak trips. This suggests that future work can assess whether more granular features can be created from the ODX time series data (e.g. indicators for the most common off-peak hours during which a passenger takes a trip) or new data sources can be introduced to more clearly delineate behavioral patterns among these categories of riders.

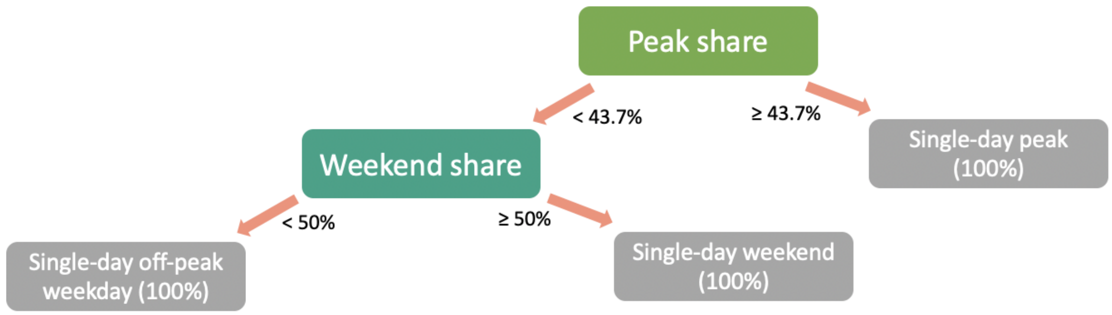


Figure 4-16: OCT for single-day riders using clustering features

4.5.2 OCT for single-day riders

The OCT for the single-day riders was as straightforward as the clustering, and robust to hyperparameter tuning. From the feature selection process, share of peak trips and share of weekend trips had been chosen as the only features. Thinking about this in the context of the multi-day OCT results in Figure 4-15, the choice of these two variables seem to be logical. The vast majority (86%) of single-day riders do not use rail at all, and by definition the active range of single-day riders is only one day (Figure 4-8. Thus the first two layers of the multi-day rider OCT diagram are not applicable here and greater emphasis is put on the timing of the journeys taken.

This simple feature set delivered a fairly clean divide of single-day riders into three clusters—36% into peak users, 41% into weekday off-peak users, and 23% into weekend users. Figure 4-16 illustrates that the first branching point is based upon the share of peak trips. Among those with less peak-period travel (that is, $\leq 43.7\%$ of journeys begin during peak hours), there is one last branching point based on weekend share that separates the weekend and weekday off-peak passengers.

The OCT exercise provides us roadmaps for what passenger ridership features were key to identifying distinctive patterns of transit usage behavior, and which rank as the first and most major partitions. It confirmed that transit mode is a key differentiator of passenger behavior, followed by the intensity of transit usage. Further, it confirmed the wide behavioral variation observed among rail rider clusters and the role of journey timing in separating these clusters.

4.6 Cluster profiling

The baseline clustering results partitioned MBTA passengers by temporal and modal ridership features built from ODX data. In this section, we leverage additional data from ODX and other sources to deepen the profile of each cluster, with the goal of enriching discussions of how our behavioral clusters relate to MBTA operations and revenue, and how they may be associated with socio-demographic patterns.

In addition to trip information, CharlieCards and CharlieTickets contain metadata on passenger fare types. These include markers for full adult fare, one of MBTA's reduced fare programs, and corporate-sponsored "Perq" passes. Separately, the MBTA also collects passenger socio-demographic information in its System-wide Passenger Survey, the latest of which was run in 2015-2017. While the methodology of the survey and its timing limit direct, granular compatibility with our analysis based on automated data sources, we still leverage station- and route-level information from this survey to gain higher-level perspective on cluster socio-demographics and cluster members' access to alternative non-transit modes. These characteristics are essential to keep in mind when we examine the COVID-19 impact on passenger mobility and policy responses in the next two chapters.

4.6.1 Cluster profiling with smart card data: MBTA operations and revenue lens

AFC and ODX data capture journeys but do not offer direct revenue figures. We examine the journey contribution of each cluster to assess the cluster's role in driving travel demand on the MBTA network. We also examine the major fare types used in each cluster; in this analysis we include Perq, the corporate pass program that made up one-third of MBTA's fare revenues, which were in turn 33% of operating revenues in 2019 [66]. The extent of MBTA's dependence on Perq income means that clusters with greater Perq presence are essential for stable operating revenues, and that the evolution of these clusters will also in part be affected by corporate negotiations with the MBTA regarding the structure and pricing point of Perq products.

This section also examines each cluster's typical weekday and weekend temporal profiles, as well as its most common origin locations. These perspectives speak to the diverging scheduling demands and geographic distribution of riders following distinct behavioral patterns, which will have implications for service adjustments during COVID-19 and the recovery.

Ridership volumes by cluster

The size of each behavioral cluster by passenger count can differ significantly from its size in terms of the volume of journeys taken. Figure 4-17 plots the clusters ranked by their journey volume against their contributions to the total number of passengers during the baseline period. From an operational perspective, the volume of journeys is the more essential metric of any behavioral cluster's systems impact.

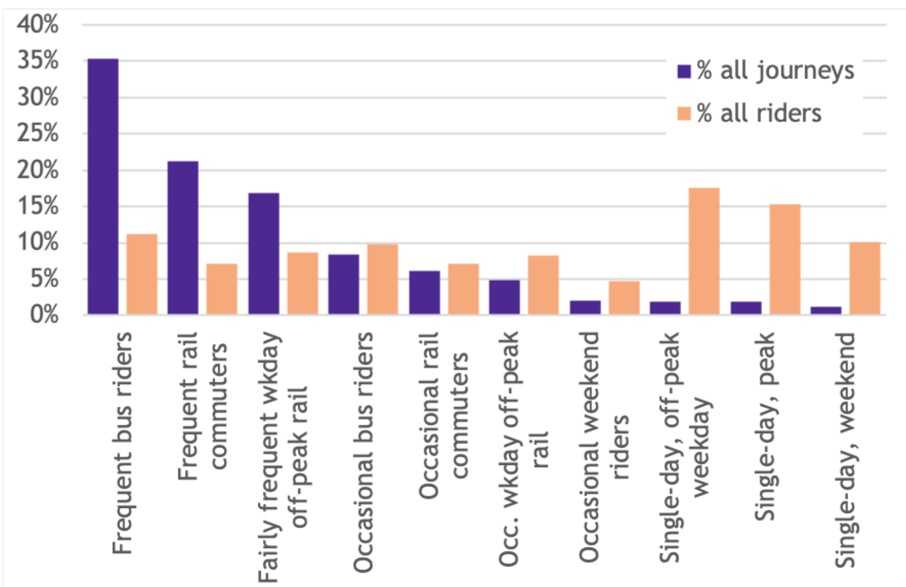


Figure 4-17: Distribution of journeys versus riders over behavioral clusters

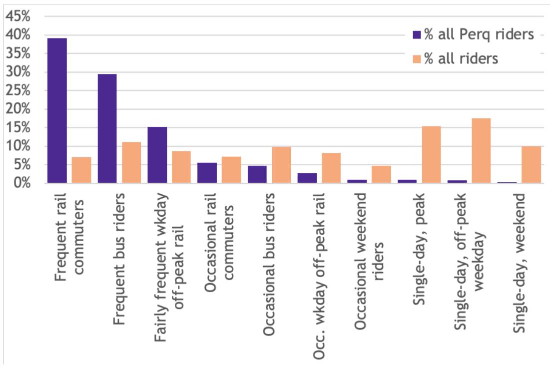
Among multi-day rider clusters, frequent bus riders contribute 35% of all journeys taken during the baseline period, far outstripping the next two clusters which are frequent rail commuters (21%) and fairly frequent weekday off-peak rail riders (17%). These also have the largest number of passengers with abnormally high journey counts, i.e. long upper tails in Figure 4-2b of up to 130 journeys over the 5-week baseline period. The frequent bus rider and frequent rail commuter clusters contribute three times more to total baseline period journeys than they do to the number of total passengers; the fairly frequent weekday off-peak riders cluster contributes two times more. Other than these three groups, each multi-day cluster contribute more to the passenger count than to the journey count.

The transit needs of passengers in these three high-volume clusters are also distinct, with one heavily utilizing buses throughout any given weekday, one riding trains only during rush hour, and one riding trains in off-peak times. Therefore, from the perspective of planning operations to meet expected travel volumes, tracking how these high-volume clusters evolve during the pandemic and adjusting to shifts in their demand patterns will be essential for understanding where services can be cut due to a drop-off in previously intensive users, and where services must be maintained due to continued heavy usage.

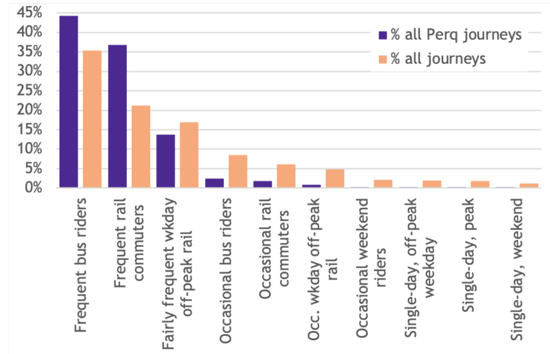
Finally, though single-day off-peak weekday riders and single-day peak riders are the largest contributors to passenger count (15% and 18% respectively) they are among the smallest contributors to journeys (1% each). Single-day weekend riders are 10% of riders and only 0.4% of journeys. Therefore, any potential pandemic-induced drop in ridership among these single-day clusters would have limited impact on MBTA operations and the revenue derived from fare sales.

Perq usage by cluster

The distribution of Perq riders across behavioral clusters is even more concentrated than the journey count. Figure 4-18a shows that frequent rail commuters dominate the pool of Perq riders (39%), followed by frequent bus riders (21%) and fairly frequent weekday off-peak riders (17%). Beyond these, the remaining seven clusters made at most modest contributions to Perq, even in cases when they make up a large share of the entire passenger pool (e.g., all single-day clusters, occasional bus riders who are 10% of all riders but only 5% of Perq ones). Given that Perq is a pass program, it is not surprising to find that only the clusters that display higher transit usage intensity are the most represented; for more occasional riders passes are less likely to



(a) Share of Perq riders versus share of all riders by cluster



(b) Share of Perq journeys versus share of all journeys by cluster

Figure 4-18: Cluster contributions to Perq corporate pass program

be cost-effective.

What is more surprising in Figure 4-18a is how dominant frequent rail commuters are among Perq riders compared to their share of the total passenger pool. This cluster contributes only 7% of all passengers but 39% of Perq riders, or 5.5 times more than a baseline scenario where each cluster contributes evenly (1:1) to Perq. Meanwhile, frequent bus riders are 11% of all riders and contribute 29% of Perq riders, or 2.6 times more. Fairly frequent weekday off-peak riders contribute 9% of all riders and 17% of Perq riders, or 1.7 times more. For occasional rail commuters this ratio is nearly 1, and all other groups contribute less to Perq than to the overall passenger count. This means that frequent rail commuters are disproportionately affiliated with the corporate pass fare revenue stream.

From a Perq journey count perspective, the disproportionate nature of frequent rail commuters' Perq subscription emerges as well. Frequent rail commuters are 21% of all journeys but 44% of Perq journeys; meanwhile bus commuters are 35% of all journeys and a much more proportionate 37% of Perq journeys. Fairly frequent weekday off-peak rail riders are a distant third in terms of Perq contributions, covering 15% of all Perq riders and a proportionate 14% of all Perq journeys. Frequent rail riders' ranking as number two in Perq journey count aligns with its position as the second largest in overall journey count behind frequent bus riders. However, a much bigger share of those journeys are affiliated with Perq-subscribing passengers than is seen in any other cluster. This again suggests the frequent rail commuters group may have disproportionate impact on this source of corporate-based fare revenue stream were their ridership behavior to shift during the pandemic or the subsequent recovery. Perq is a pass program, so from a revenue perspective it is the rider numbers that count more than the journeys. Nevertheless, how passengers' journey counts evolved during COVID-19 and the recovery will affect the number of passengers from each cluster who still find Perq a cost-effective program at current pricing points.

Lastly, we note that the top Perq sponsor companies by cluster are fairly similar across the three clusters with significant Perq subscriber shares. Though our data does not include the names of the companies themselves due to confidentiality agreements, we are given that the most frequently recurring sectors among these clusters are hospitals and universities. Two third-party corporate benefits providers that serve multiple employers—WageWorks and Edenred—are also among the most frequently seen, but we cannot tie these to specific companies. These organizations are generally located in the downtown Boston or Cambridge areas. The top origin locations in the Perq-heavy frequent rail commuters cluster are in heavily commercial downtown areas, and areas with university and medical facilities. The situation is similar in the fairly frequent weekday off-peak rail cluster.

Cluster temporal profiles

The OCT exercise and feature selection process both noted the importance of the time of day of ridership for distinguishing between clusters. Differences in the timing for transit demand between major rider groups is a central consideration for transit planners, for reducing crowding and providing service to meet demand patterns.

In this subsection, we take a more granular view at temporal patterns of clusters beyond the simplified summary features used for the actual clustering. Figure 4-19 plots the distribution of journeys taken by riders by time of day on weekdays, for each multi-day rider cluster as well as for all multi-day riders. Because the MBTA is a tap-in only system, all time of day values indicate the beginning of journeys. The y-axes of the plots show the percentage of journeys in each cluster that began

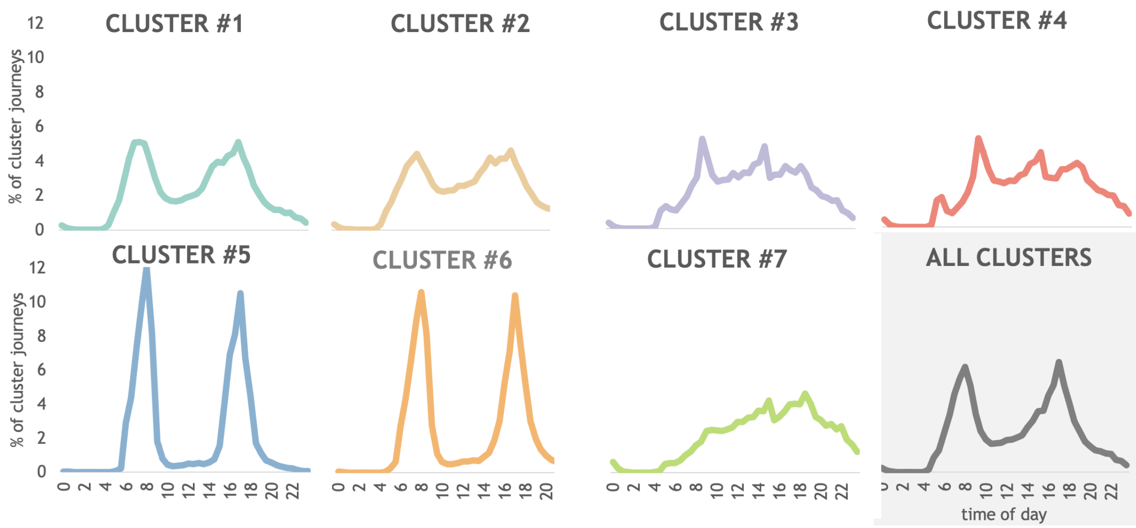


Figure 4-19: Average weekday temporal profiles by multi-day cluster, and for all baseline journeys

during the time of day specified on the x-axis. Time-of-day of travel is aggregated to half-hour increments to reduce noise.

During the weekdays, the frequent and occasional rail commuter clusters capture the peak-oriented travelers as expected. Figure 4-19 makes clear that these two commuter groups are the ones driving the peak demand visible when we aggregate all multi-day riders, and that the peak-period travel for the frequent and occasional bus riders are much less pronounced. The ratio of peak travel at 8 AM to mid-day at 11 AM is 25.5 for occasional rail commuters (cluster 5) and 15.2 for frequent rail commuters (cluster 6); both groups had almost no travel volume during the middle of the day. By contrast, this ratio is only 2.7 for frequent bus riders and 1.9 for occasional bus riders, who show much more sustained demand throughout the day.

The fairly frequent weekday off-peak rail rider cluster still shows some presence of morning and evening peaks despite a much more robust midday volume. This cluster's evening ridership also tails off much more gradually later into the night than do the bus clusters or rail commuter clusters, potentially indicating recreation, errands, or evening/late shift work. Lastly, weekend rail riders (cluster 7) do have some weekday volumes despite its orientation towards weekend demand; the weekday travel was mostly in the middle of the day trending up for the evening. This could

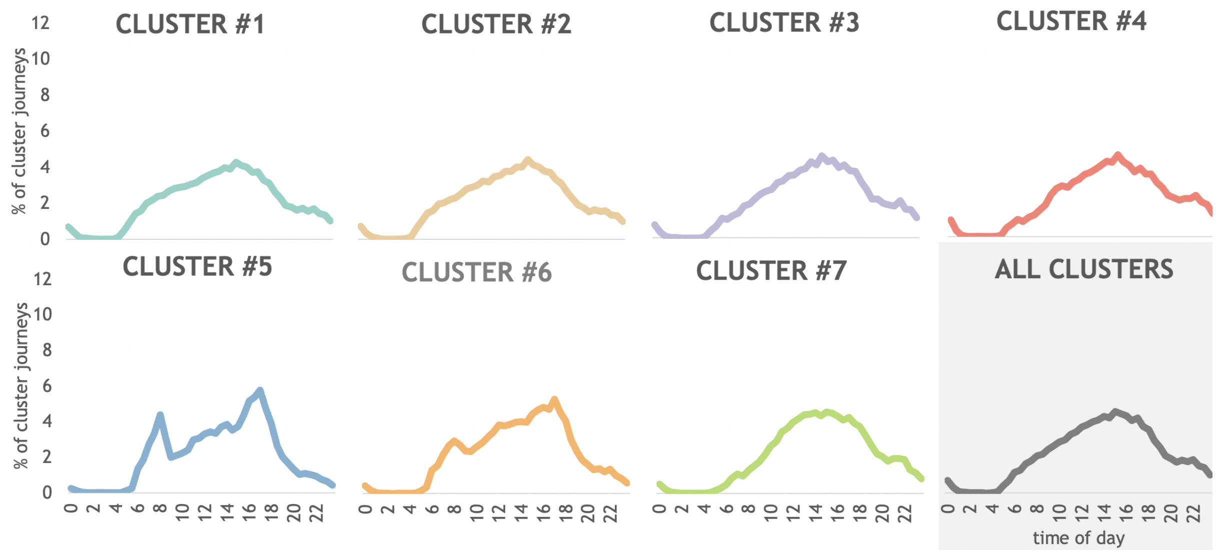


Figure 4-20: Weekend temporal profiles by multi-day cluster, and for all baseline journeys

again reflect the leisure-oriented nature of this cluster's transit use.

Similarly, Figure 4-20 plots the average temporal profile for a weekend day, where we see much less cluster-level variation in the distribution of journeys over time of day. All clusters reached maxima during the afternoon and evening periods of the weekend, with the weekend rail riders (cluster 7) showing slightly more sustained volumes in the midday. The two rail commuter clusters (clusters 5 and 6) surprisingly still show some morning peak, especially for occasional rail commuters. This could indicate some weekend work travel for those with non-traditional work schedules (e.g. for hospital staff) but our data does not give us the information to pinpoint the reason behind these temporal patterns. The bus rider clusters (clusters 1 and 2) have slightly higher volumes in the morning as well, but these are not as prominent as for rail commuters.

Figure 4-20 is in relative terms since the y-axis indicates the share of cluster journeys taken during a particular time. In absolute numbers, the occasional rail commuters cluster has negligible weekend travel volumes. Frequent rail commuters have substantial volumes but are still only one-third of the journey counts observed for frequent bus riders, which tally the largest weekend volumes followed by fairly frequent weekday off-peak rail riders. The weekend bumps in morning and evening travel of occasional rail commuters are therefore of little operational significance, while the similar pattern among frequent rail commuters are more impactful for weekend planning and travel demand.

Cluster spatial profiles

By definition, transit operations require spatial analysis. This section provides a high-level look at the distinctive spatial profiles of each multi-day and single-day cluster which provides further reason for MBTA planning to separately consider the mobility needs of bus- versus rail-oriented riders.

Figure 4-21 presents a heat-map table of the top ten most frequented origin stations for each behavioral cluster. Highlighted in green are the stations or bus stops that are particular to only one or two clusters, with no representation among the majority of other segments.

	Frequent bus riders	Occasional rail commuters	Occasional weekend riders	Frequent rail commuters	Occasional bus riders	Fairly frequent wkday off-peak rail	Occ. wkday off-peak rail	Single-day, peak	Single-day, off-peak weekday	Single-day, weekend
Rank	**Cluster 1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
#1	Forest Hills	South Station	Harvard	South Station	Forest Hills	Park Street	South Station	North Station	South Station	Harvard
#2	Forest Hills (bus, lower deck)	North Station	Park Street	Downtown Crossing	Nubian (bus)	Downtown Crossing	North Station	South Station	North Station	Park Street
#3	Harvard	Back Bay	South Station	Kendall Square	Forest Hills (bus, lower deck)	Central Square	Downtown Crossing	Downtown Crossing	Downtown Crossing	South Station
#4	Nubian (bus)	Kendall Square	Central Square	Park Street	Forest Hills (bus, upper deck)	Harvard	Harvard	Back Bay	Harvard	North Station
#5	Forest Hills (bus, upper deck)	Downtown Crossing	Kenmore Square	State Street	Harvard	Kendall Square	Park Street	Park Street	Park Street	Boylston
#6	Downtown Crossing	Porter Square	North Station	Central Square	Downtown Crossing	South Station	Kendall Square	Harvard	Back Bay	Kenmore Square
#7	Ashmont (bus)	State Street	Kendall Square	North Station	Haymarket	Maverick	Back Bay	Alewife	Haymarket	Government Center
#8	Central Square	Central Square	Boylston	Alewife	Haymarket (bus)	North Station	Copley Square	Kendall Square	Copley Square	Hynes
#9	Sullivan Square	Harvard	Hynes	Back Bay	Ashmont (bus)	Copley Square	Central Square	State Street	Charles MGH	Central Square
#10	Charlestown Garage	Charles MGH	Airport	Davis Square	South Station	Airport	Maverick	Copley Square	Kendall Square	Prudential

Figure 4-21: Top 10 origin stations by cluster

Larger intermodal stations that serve as transfer hubs unsurprisingly top the list for multiple clusters. For example, South Station appears for nine out of 10 clusters, Downtown Crossing for eight, North Station for eight, and Park Street for seven. However, more rail-dependent clusters do not list any bus stops among their top origins, reflecting the less intermodal nature of the rail-oriented passenger clusters that we first observed in Figure 4-6. Meanwhile, the frequent bus riders cluster lists four bus and six rail entry points among its top 10 and the occasional bus rider cluster's top origins were split in half between bus and rail.

ODX modal data also support this observation that bus riders are more likely to use a combination of bus and rail, compared to rail-oriented rider clusters. Only 50% of both the frequent bus riders and the occasional bus riders spend the overwhelming majority of their trip stages traveling by bus (where overwhelming majority is defined as over 75%). The other half of the passengers in these clusters use a well-mixed combination of rail and bus.

Meanwhile, across all multi-day rail clusters the share of passengers with an overwhelming majority (75% or more) of rail-based trip stages ranged from a low of 71% for occasional weekend riders to a high of 86% for occasional weekday off-peak rail riders. Single-day clusters are also heavily dominated by rail, and because of the small average number of journeys per cluster (just under 1.5 for each of these groups) the share of passengers in each single-day rider cluster taking the overwhelming majority of their trip stages by rail was generally high, ranging from 83% among the off-peak weekday cluster to 87% for the weekend users. The off-peak weekday cluster contains the largest number of single-day passengers dependent on bus. Supporting the rail network thus appears essential for meeting the transit needs of both rail- and bus-oriented repeat riders as well as single-day visitors, though the specifics would require extensive spatial and network analysis.

There were also geographic differences in popular origins across clusters. While the rail clusters generally favor stations in downtown Boston or Cambridge, the top stops for both frequent and occasional bus riders are heavily tilted towards South Boston locations including Forest Hills, Ashmont, and Nubian. These stations are intermodal with train access but still are not popular among rail-oriented clusters. The frequent bus riders' top stations further include the northern Boston Sullivan Square and Charlestown area, which is known as an intermodal transit hub connecting bus, subway, and commuter rail. These are also not frequent origin points for rail riders. Figure 4-22 maps the top origins of frequent bus riders (cluster 1), which is representative of the occasional bus riders as well, against the top origins of frequent

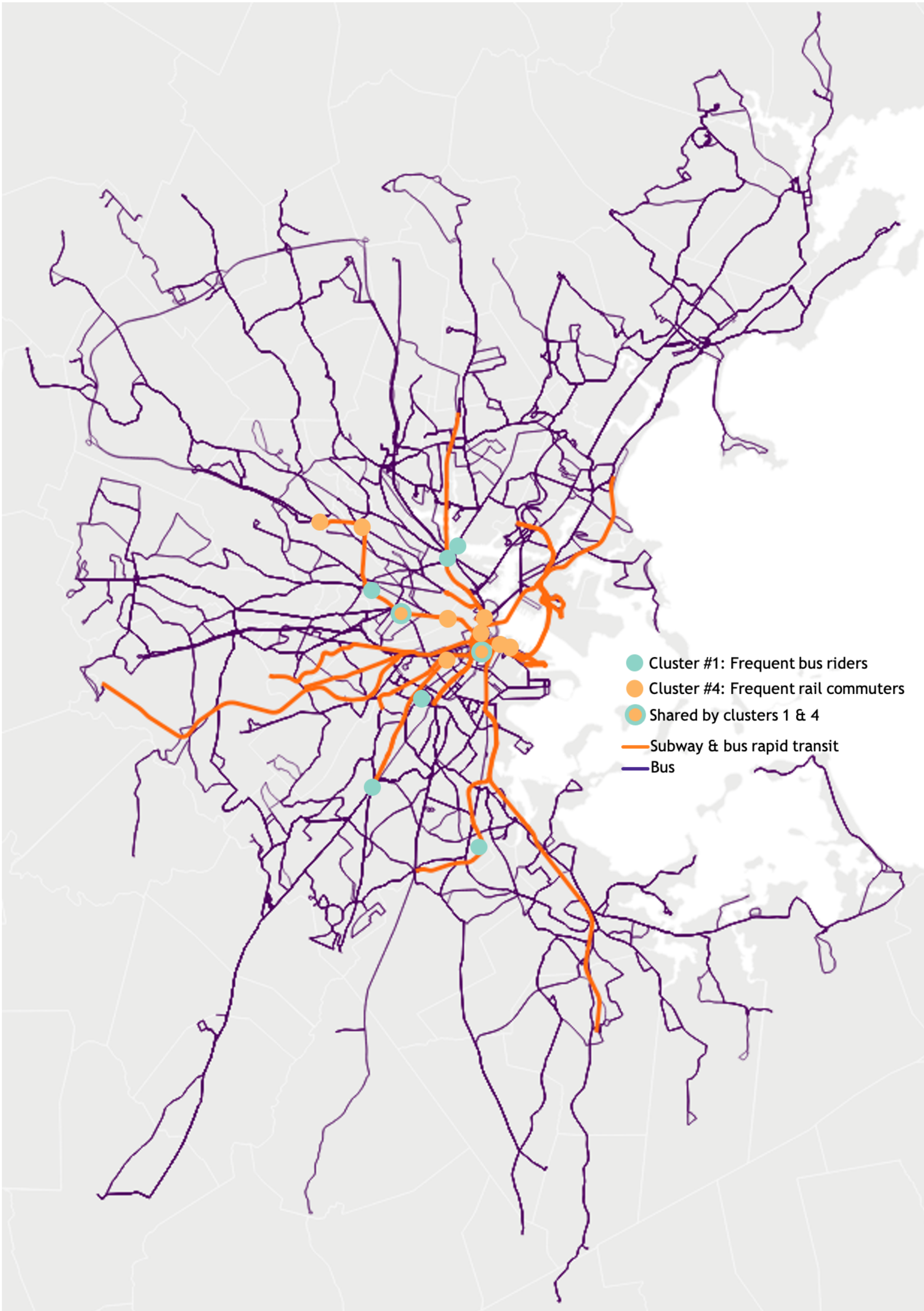


Figure 4-22: Map comparing top origin stations for Cluster #1 frequent bus riders v. Cluster #4 frequent rail commuters

rail commuters (cluster 4) which is fairly representative of other rail clusters. From this the spatial separation between popular stations used by the bus and rail clusters becomes clear, with Central Square and Downtown Crossing being the two heavily used by both subgroups. For rail commuters, seven of the ten top stations are specifically along the Red Line. This already begins to suggest geographic and associated socio-demographic differences between clusters that can affect MBTA's service cut decisions once COVID-19's uneven impact on churn is felt across clusters.

Among rail-oriented clusters, the off-peak riders stand out. Both fairly frequent weekday off-peak rail riders and occasional weekday off-peak rail riders heavily use Maverick station, which is intermodal between several bus lines and the Blue Line. It is also located in East Boston, away from the central Boston and Cambridge areas highlighted by other rail clusters. The fairly frequent weekday off-peak rail group includes the airport stop as well among its top ten origins, which is proximate to the Maverick station in East Boston.

Journey tap-in times and origin data within ODX and AFC highlight the temporal and/or geographic separation of bus riders from rail riders, and between the fairly frequent weekday off-peak riders versus other rail clusters.

4.6.2 Cluster profiling with smart card data

Usage of pay-as-you-go versus transit passes

The ODX data used for clustering also contains smart card metadata on fare product types that can be used to partially characterize the socio-demographic composition of riders within each behavioral cluster. The fare product types recorded in AFC include the 1-day pass, 7-day pass, monthly pass, and pay-as-you-go; Perq corporate pass usage was also separately available from the Perq database as previously discussed.

The differentiation between pass types is of interest for supporting socio-economically vulnerable riders in Metro Boston, because the higher up-front cost of longer-term fare passes can prevent even riders making more frequent trips from adopting them [41]. A one-day pass, for example, costs $12.75/day in 2020, while a 7-day pass is only

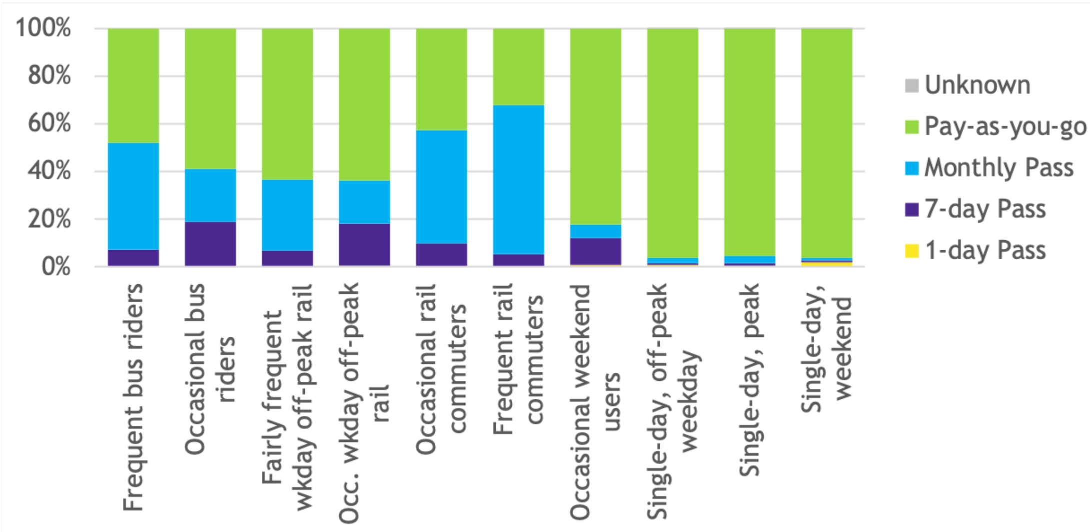


Figure 4-23: Fare product type distribution by cluster

76% more at $22.50 for six additional days, and a monthly Link pass is $90 for one calendar month (there is also a bus-only monthly pass for $55 per calendar month, but our data does not distinguish this). Further, corporate-subsidized Perq passes are tied to company sponsorship, making this pass type inaccessible for passengers not employed by companies large enough to negotiate and design such contracts with the MBTA.

Pairing clustering results with this smart card meta-data shows extremely low 1-day pass usage across all clusters. The single-day weekend users, who are likely leisure-oriented, have the highest uptake of one-day passes at 1.8%. For the 7-day pass, the uptake rate was highest among occasional bus riders where 19% typically tapped in with such passes. The occasional weekday off-peak rail cluster is a close second at 17.7%. Monthly passes are the most popular pass type across all clusters except for the leisure-oriented occasional weekend users and single-day weekend users. Frequent rail commuters are most likely to take advantage of monthly passes (62%), with occasional rail commuters (48%) and frequent bus riders (45%) in distant second and third place.

It is notable that occasional rail commuters, despite a significantly lower system usage, still had a roughly equivalent share of monthly pass adopters as frequent bus riders. For occasional rail commuters, Figure 4-7 shows that the interquartile range

(IQR) fell between 5 and 17 days for active range and 3 to 10 days for active days. By contrast in the frequent bus riders cluster, nearly all riders exhibited an active range of 30 or more days while the IQR for active range was 13 to 25 days. Part of the reason for occasional rail commuters' out-sized monthly pass usage compared to more transit-engaged clusters may be due to fare design: given that subway rides were \$2.40 per trip by CharlieCard (\$2.75 by CharlieTicket) versus \$1.70 for bus, fewer trips are needed to justify monthly pass subscriptions by rail-oriented rather than bus-oriented users.

However, comparing the monthly pass subscription rate of occasional rail commuters against that of fairly frequent weekday off-peak riders suggests that this is not the entire reason. Fairly frequent weekday off-peak riders have greater transit engagement than occasional rail commuters, with an IQR of 26-33 days for active range and 6-17 days for active days. Both clusters typically ride rail and pay rail fare prices. Yet, the off-peak cluster's monthly pass subscription rate is only 29.9%, and it shows high reliance on pay-as-you-go at 63% which is more similar to occasional bus and occasional weekday off-peak riders than to frequent usage clusters. Occasional rail commuters' larger pass subscription rate may reflect a greater capacity for affording the convenience of a monthly pass; it also may be because fairly frequent weekday off-peak rail riders are more likely to have reduced fares (e.g. through senior cards) which lowers the purchase price at which a monthly pass becomes the most cost-effective choice. Figure 4-25 below does show that this cluster is one of the largest for reduced-fare customers which seems to align with these hypotheses, but our data does not allow us to draw causal conclusions.

Frequent rail riders are also the most likely to have monthly passes paid through the corporate Perq program. Figure 4-24 illustrates that 33% of this cluster's members had Perq sponsorship, with frequent bus riders a far distance second at 16% followed by fairly frequent weekday off-peak rail riders at only 10%. Note that this graph differs from the preceding Figure 4-24 in that it depicts Perq uptake rate *within each cluster*, rather than assessing how much each cluster contributes to the total Perq corporate adopter pool. The fact that Perq sponsorship is twice as high among frequent rail

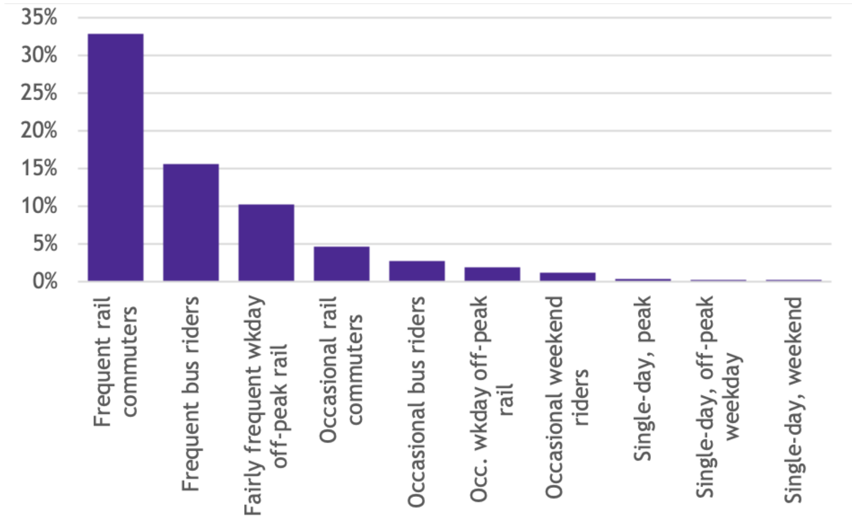


Figure 4-24: Share of Perq users by cluster

commuters than among frequent bus riders and fairly frequent weekday off-peak rail riders, who are also intensive users of the system, again suggests structural differences between clusters in terms of the type of people riding and their access to corporate-negotiated and corporate-subsidized transit benefits packages.

Full and reduced fare user types

Smart card meta-data carries "user type" information indicating whether the card makes adult full-fare transactions (either as a pass validation or a pay-as-you-go trip) or reduced fare transactions. This give a direct window into the transit usage behavior of the socio-economically vulnerable who qualify for smart cards with adjusted fares. The main fare categories are detailed in the list below. Using a CharlieCard, full fares were \$1.70/trip for bus and \$2.40/trip for rail one way. At reduced fares, bus was \$0.85/trip and rail \$1.10/trip. During the baseline 2020 period, CharlieTicket full fares were more expensive for rail at \$2.75, but this uptick for paper ticket versus CharlieCard fares was eliminated during COVID-19.

1. **Adult**: Typical full-price fares.

2. **Blind**: Free MBTA pass for legally blind riders and accompanying guides.

3. **RIDE**: Reduced fare pass for those with disabilities who have difficulties accessing the bus or rail (subway, trolley) services. RIDE passengers can access paratransit services booked on-demand, and generally use RIDE more for paratransit services rather than typical transit services.

4. **Senior**: Reduced MBTA fares for senior citizens aged 65 and over.

5. **Student**: Various tiers of discounts offered to students, from middle school through university. The program is offered through institutions to their affiliated students.

6. **TAP**: TAP offers reduced MBTA fares for people with temporary or permanent disabilities and Medicare cardholders. These riders typically use the main MBTA system, unlike RIDE passholders who rely mostly on paratransit.

7. **Other**: This group includes all other groups, such as other types of youth passes, police, firefighters, public officials, and retired MBTA employees. It also counts short fares, which indicates when a driver allows a passenger to board a vehicle using a card without enough value to cover the full fare.

Notably, we include short fares in the "other" category. Short fares mostly occur on buses and overground segments of the Green and Mattapan lines since station faregates on subway sections of the rail network do not allow conductor or driver discretion in granting entry to prospective passengers paying partial fares. Thus, the occurrence of short fares in the data will naturally occur mostly among the two bus-oriented behavioral clusters simply due to the nature of this fare label. To avoid drawing attention to this artificial pattern, we decided not to separate out this category from the other smaller fare types.

Further, the data does not trace fares paid by cash since those transactions are lumped under one "card number" in AFC and thus not traceable through smart card data. This is a significant limitation to the ability of this ODX-based analysis to track vulnerable transit riders, but one that cannot be addressed with the data sources on hand for this thesis. Follow-up studies can use survey-based methods to assess transit usage patterns and trip purposes for these riders which is key for equity in planning.

Figures 4-25 decomposes each behavioral cluster by user type while 4-26 flips the analysis, decomposing each user type by cluster to illustrate which clusters contribute the most to vulnerable rider categories. Figure 4-25 shows that adult fare riders compose the majority of any cluster, ranging from a low of 69% for frequent bus riders up to a high of 96.8% for single-day weekend riders. Students, seniors, TAP users, and Other are the remaining major user types appearing in multiple clusters; the prevalence of Other is driven mostly by short fares among frequent bus riders.

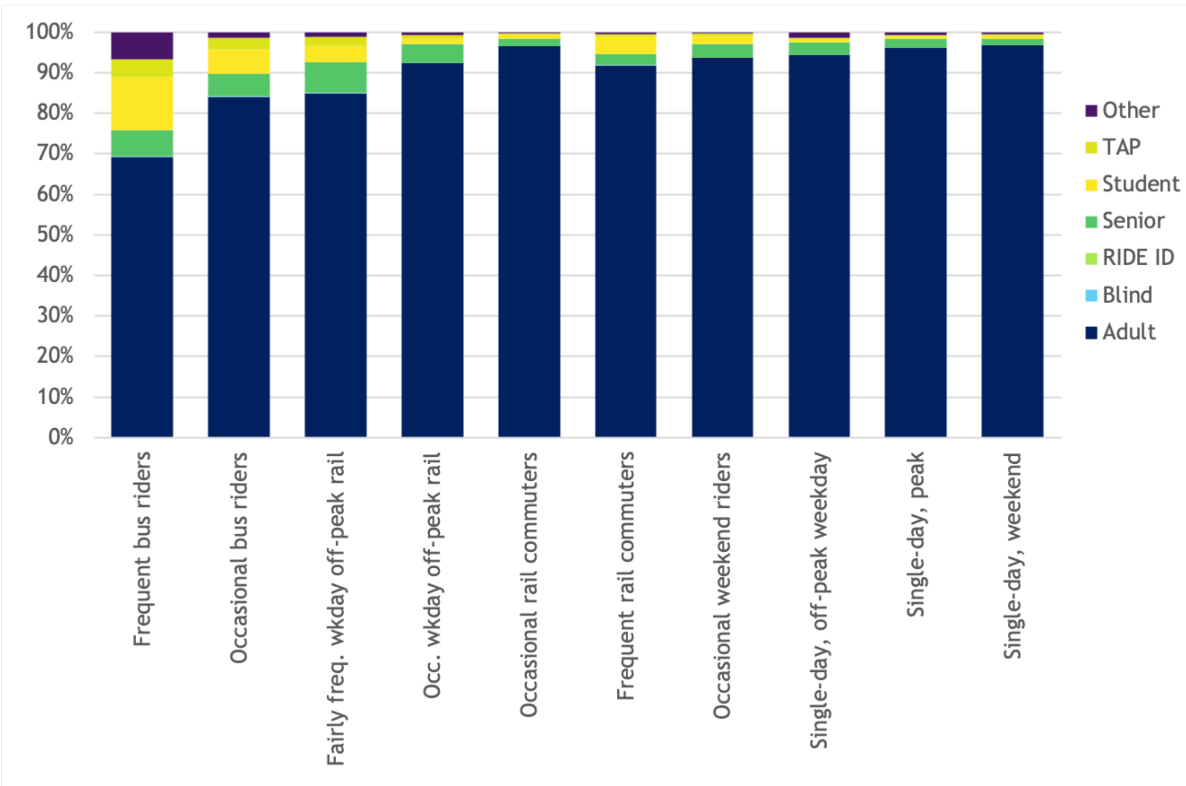


Figure 4-25: Passenger behavioral clusters decomposed by smart card user type

Frequent bus riders, occasional bus riders, and fairly frequent weekday off-peak rail riders have the largest share of passengers with discounted or free fares. Only 69% of frequent bus riders, the largest cluster, are adult full-fare. 13% are students, 6.4% senior reduced fare riders, 4.4% TAP riders, and 6.6% Other composing mainly of short fares (5.6%). There is a greater portion of students here than anywhere else,

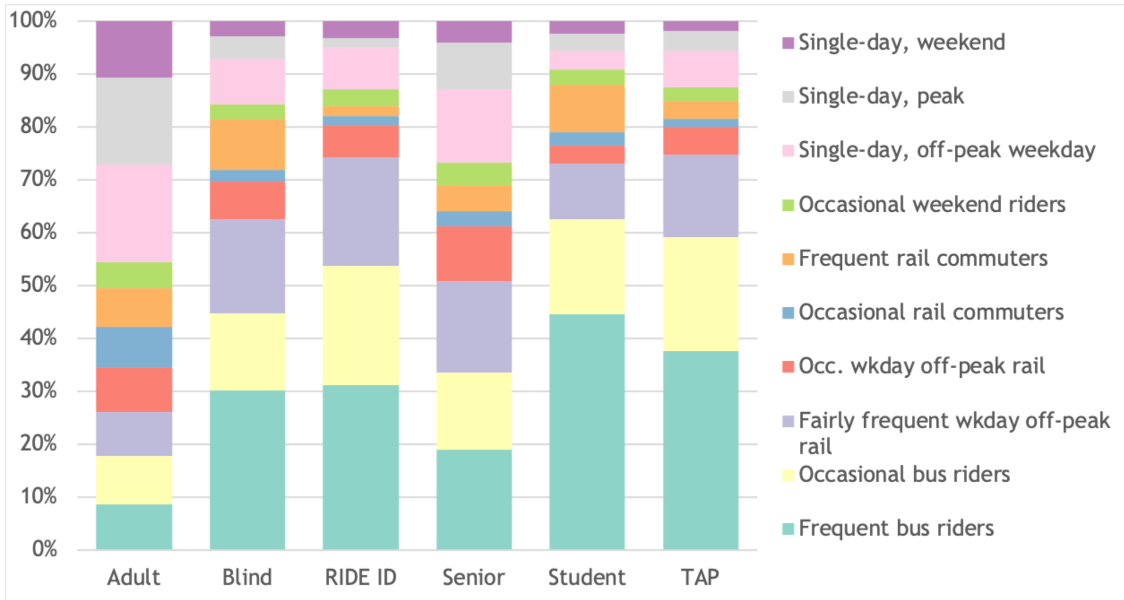


Figure 4-26: Smart card user types distribution across clusters

which also translates to the greatest absolute number of students since frequent bus riders is the largest cluster. Occasional bus riders are 84.1% adult full-fare riders, 5.6% senior pass holders, 6% students, 2.9% tap, and 1.3% Other (mostly youth fares). The two bus-oriented clusters therefore are also the categories with the largest share of reduced fare riders, with seniors and students being the most notable categories. The TAP portion within frequent bus riders is also higher than for any other cluster.

The fairly frequent weekday off-peak rail rider cluster has a larger share of senior passengers than any other cluster (7.5%); it is also 3.9% students and 2.3% TAP riders. Adult riders only make up 84.9% of the cluster. The occasional weekday off-peak rail group also has a notable senior share (4.7%). Finally, while single-day clusters were dominated by adult fares, seniors were the only reduced-fare category with a notable presence, again preferring weekday off-peak travel (3%). Together, these suggest that services meant to target seniors need to be conscientious of planning for off-peak weekday travel demand.

Figure 4-26 also highlights the uneven spread of reduced fare riders across the behavioral clusters, again pointing to diverging and distinct patterns of transit use among the socioeconomically vulnerable versus other transit riders. The first column, adult fare users, is dominated by single-day clusters each of which contributes 11-19% of all riders in this user type. The rest of the adult fare user type is spread fairly evenly among multi-day clusters (7-9% each) except for occasional weekend riders which was the smallest contributor (5%). This already suggests that the larger clusters, such as frequent bus riders, occasional bus riders, and fairly frequent weekday off-peak rail riders are under-represented in the pool of full-fare riders compared to their bigger overall size.

Moving to the right in Figure 4-26, the reduced or free fare user types are again clearly concentrated among the two clusters of bus riders as well as fairly frequent weekday off-peak rail riders. The student, TAP, RIDE, and Blind user types are especially heavy on passengers from the frequent bus rider cluster, which makes up 44.5%, 37.6%, 31.2% and 30.1% of those reduced fare user types, respectively. The occasional bus riders group is another major contributor, making up 22.5% of RIDE, 21.6% of TAP, 18.0% of student, 14.6% of senior, and 14.6% of the blind user types. Fairly frequent weekday off-peak rail cluster is a similarly important contributor to vulnerable user types, making up 20.5% of RIDE, 17.9% of blind, 17.3% of senior, 15.5% of TAP, and 10.5% of student user types. Among seniors, these three clusters are the top contributors. Additionally, single-day off-peak weekday riders also made up 13.9% and occasional weekday off-peak rail riders 10.2%. This again emphasizes seniors' propensity to use the public transit system during off-peak weekday periods.

In summary, Figure 4-26 confirms that socioeconomically more vulnerable riders using reduced fares have a greater tendency for bus-based travel and off-peak weekday travel. The temporal and modal separation between these riders and the majority of full adult-fare riders again highlight the distinctive ways in which sub-populations within Metro Boston separately engage with the transit system.

Passenger-level spatial distribution of origin stations

A previous section discussed the greater representation of South Boston origin locations for journeys undertaken by passengers within bus-oriented clusters. For the frequent bus riders cluster, the most frequent origin station is the intermodal Forest

Hills station in southern Jamaica Plains; the subway platform and the upper and lower bus decks each registered as a top entry point for this cluster (Figure 4-21, Figure 4-22). Bus boardings at Nubian station in the Roxbury neighborhood also topped the list. The occasional bus rider group prominently features these origin locations as well. Forest Hills and Ashmont are rapidly developing neighborhoods with varied socio-economics, while Roxbury is a diverse area with 32% white, 53% Black/African American, 29% Hispanic, 3% Asian/Pacific Islander, and 4% other [67]. By comparison Boston as a whole is 45% white, 23% black/African American, 19% Hispanic, and 9% Asian/Pacific Islander.

The ODX journey origins data also supports analysis of the distribution of locations from which passengers in each cluster access transit. We calculate Shannon entropy scores for journey origins at the passenger level and plot the distribution of entropy values found in each cluster. This helps to describe the variation in origin stations used by riders in each cluster; the higher the median entropy, the more passengers in that group tend to rely on transit to access a wider variety of their mobility needs. We do this only for the multi-day rider clusters, since the single-day riders averaged under 1.5 journeys during the baseline study period which is too small a number for meaningfully calculating entropy.

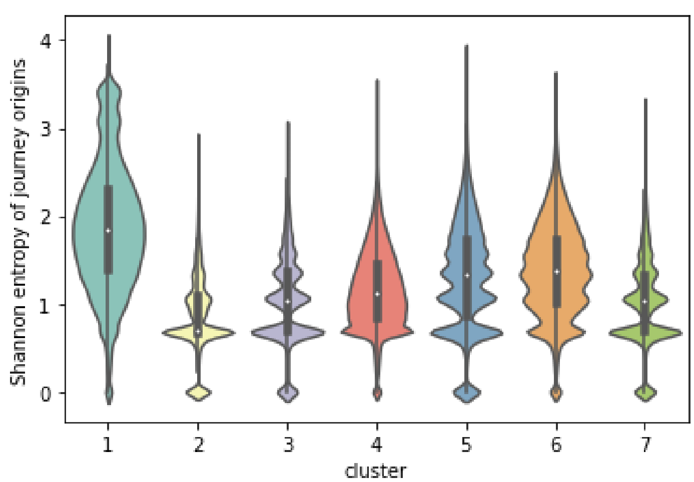


Figure 4-27: Distribution of passenger-level Shannon entropy scores of journey origins for each multi-day rider cluster

Figure 4-27 illustrates the higher level of entropy in journey origins generally found among frequent bus riders compared to all other multi-day clusters. Though the entropy among frequent bus riders also has longer tails than other clusters, this group appears to use the bus from a wider range of origins even compared to frequent rail commuters, though the distribution of total journeys per passenger was similar between the two groups (Figure 4-2b). This suggests consistent bus users may be more dependent on the bus for a wider range of mobility needs. We note, however, that there are also many more bus stops spaced closer together (8,047 stops) than there are rail and rapid transit stops (166 stops) in Metro Boston, which could contribute to greater entropy in the bus results if we assume that bus riders sometimes board using stops adjacent to their most typical ones [68]. However, it is reassuring for this analysis that occasional bus riders do not appear to outstrip any of the occasional rail rider groups in terms of journey origins entropy.

4.6.3 Cluster profiling with survey data

This sub-section leverages the MBTA's 2015-2017 System-wide Passenger Survey to give socio-demographic color to the profiles of each behavioral cluster, then to draw high-level conclusions about access to alternative modes and ease of access to transit stations by cluster. As discussed in Chapter 3.2, this survey was conducted at the station level for subway services and at the route level for bus and above-ground rail segments. Further, survey respondents were sampled from any passenger in the station or bus/trolley vehicle including those entering, transferring, or exiting; by contrast, ODX data only records passenger entry. Because of these methodological differences between the survey and ODX data, it was not possible to perfectly join each passenger card ID in the ODX database to an exact demographic profile. Additionally, because bus riders were surveyed at the less granular route level while subway riders were surveyed at the station level, we expect that the extrapolated results for the frequent and occasional bus rider clusters may look more similar to each other than for the rail-oriented clusters. Therefore, we do not use this data to draw causal inferences or granular analytical results, but only use it for high-level insights that give a sensibility check to the results from our smart card meta-data analysis.

Difficulties of socio-demographic inference to transit riders are not unique to our survey dataset. The most common alternative for socio-demographic data, the census, also faces problems of specificity when applied to AFC or ODX, because analysis based on the census seeks to infer the socio-demographics of transit riders based on the the location of tap-ins [36, 33, 35, 2, 34]. Not all individuals in a census tract where a tap-in occurs are regular transit users, and a tap-in does not necessarily indicate that the rider resides and therefore is representative of the origin census tract.

Socio-demographics

Given these caveats, Figure 4-28 estimates each cluster's income bracket breakdown based on extrapolation from the System-wide Passenger Survey. The frequent and occasional bus rider clusters outstrip all others in the share of households earning less than $76,000 annually, with 51-52% in brackets below this threshold while all other clusters contain at most 44% below this threshold. There is a significant share of riders extrapolated as preferring not to share income information based on the survey data. This share is steady at 16-17% across all clusters, but the underlying way in which this category is spread among the known income buckets may not be the same across clusters. This introduces more caveats to our income analysis. Keeping these in mind, it appears that the frequent and occasional bus rider clusters reflect the transit usage behaviors of lower income demographics. This aligns with our earlier finding that bus-based clusters are more reflective of usage patterns by reduced-fare user types including students, those paying short fares, seniors, and those with disabilities, based on the more accurate direct analysis of ODX smart card meta-data.

On the other hand, the fairly frequent weekday off-peak rail rider cluster does not appear to have as high a share of lower income bracket riders, despite earlier ODX smart card meta-data analysis showing that several key reduced-fare user types exhibited ridership behavior captured by this cluster. This can be reflective of the composition of the reduced fare user types dominant in this cluster as well as the occasional weekday off-peak rail cluster. Holders of senior reduced fare cards make up the majority of those not paying full fares within these clusters (Figure 4-25), but

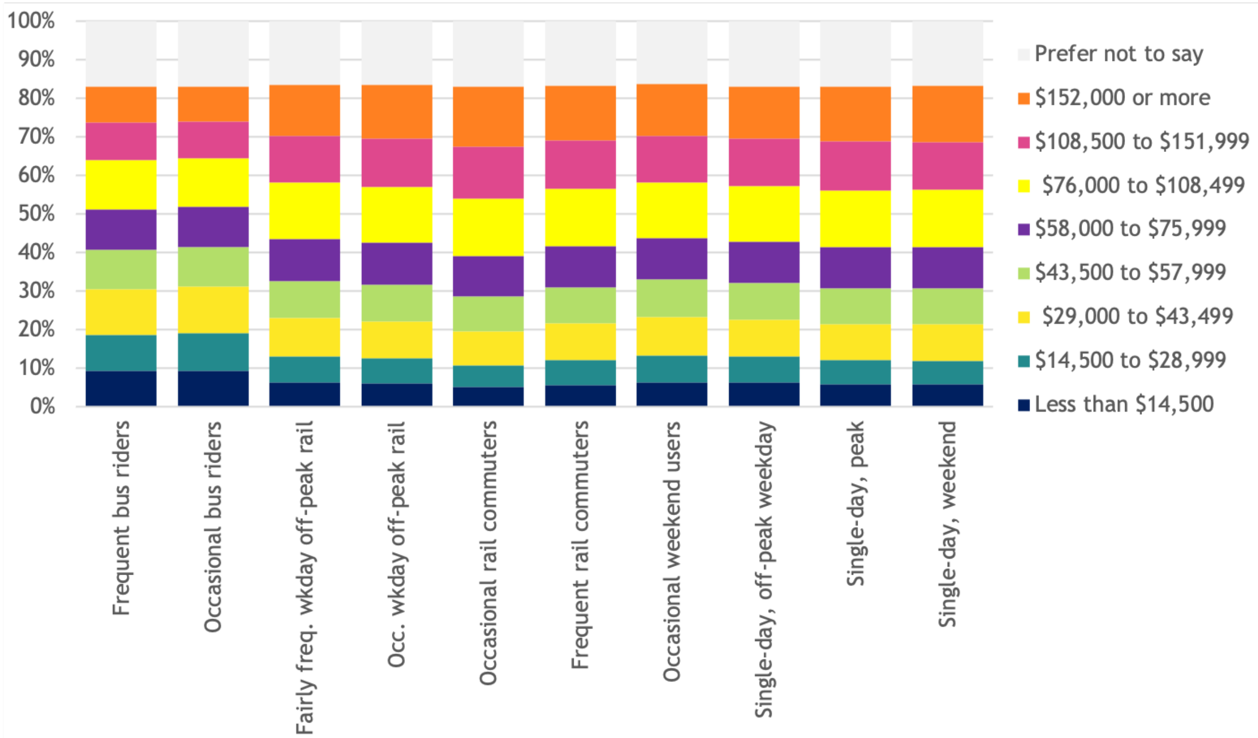


Figure 4-28: Estimated household income breakdown by cluster based on MBTA survey data

seniors may still have substantial work or investment income.

Figure 4-29 plots the extrapolation of ethnic breakdown by cluster based on survey data. The plot suggests that the frequent bus riders cluster and occasional bus riders cluster hold higher proportions of people of color, in particular black or African American riders and American Indian or Alaska Native riders. The American Indian or Alaskan Native share of each bus rider cluster was around 2%, twice that of all other clusters. Black or African American riders were around 19% of each of the two bus rider clusters, compared to 9%-11% for the other clusters. Though the System-wide Passenger Survey does not offer the passenger-level accuracy of our ODX data, it seems to align with our findings from smart card meta-data that bus-oriented riders could be from more reduced-fare groups.

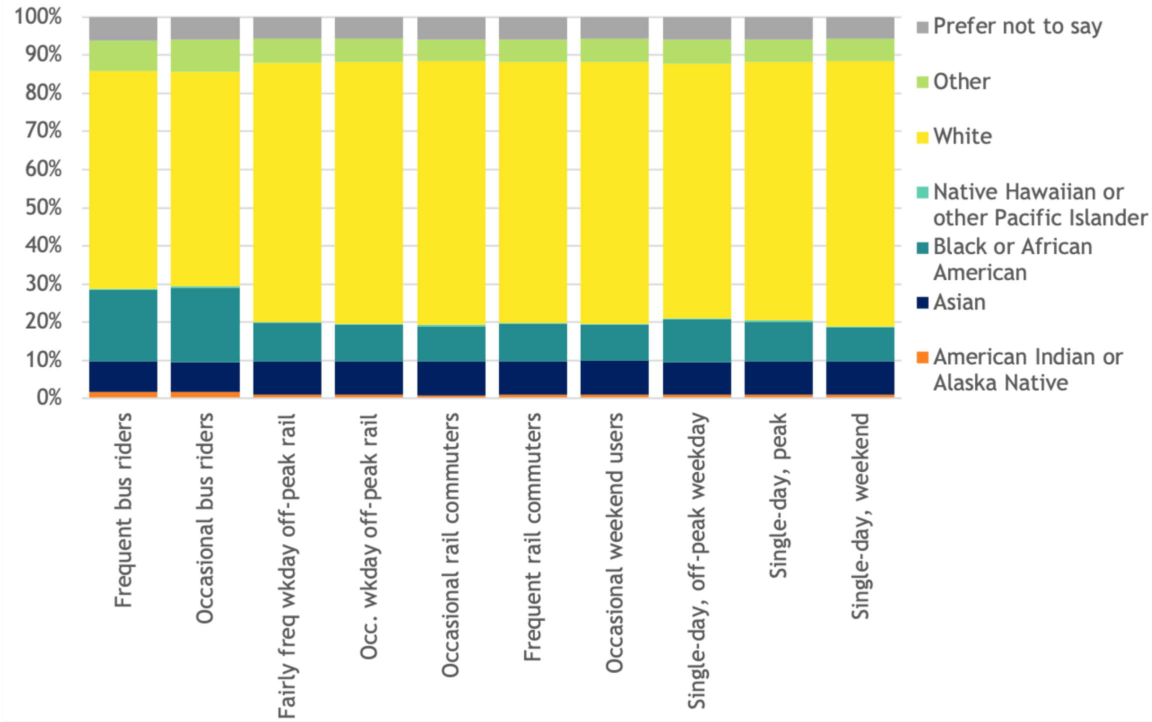


Figure 4-29: Estimated ethnicity breakdown by cluster based on MBTA survey data

Modal alternatives

The System-wide Passenger Survey includes information on access to alternative modes, number of household vehicles accessible to the transit rider, and the mode used to access the beginning of a rider's transit journey. Figure 4-30 depicts, by cluster, passengers' alternatives to transit for the journey they were conducting when interviewed for the survey. The extrapolated figures suggest frequent and occasional bus riders are the least likely to drive alone as an alternative to transit, with 14-15% stating they would drive alone compared to 18-21% in other groups. Extrapolations also suggest, however, that they are not less likely to join a carpool as an alternative. They may instead be slightly more likely to take a different MBTA transit service.

These results seem to align with patterns observable in Figure 4-31, which suggest frequent and occasional bus riders are the most likely clusters to have no usable household vehicles, as are occasional weekend riders—for each of these clusters, roughly one-third responded they had zero household vehicles. Fairly frequent weekday off-peak and occasional weekday off-peak rail riders are close behind at 30-31%.

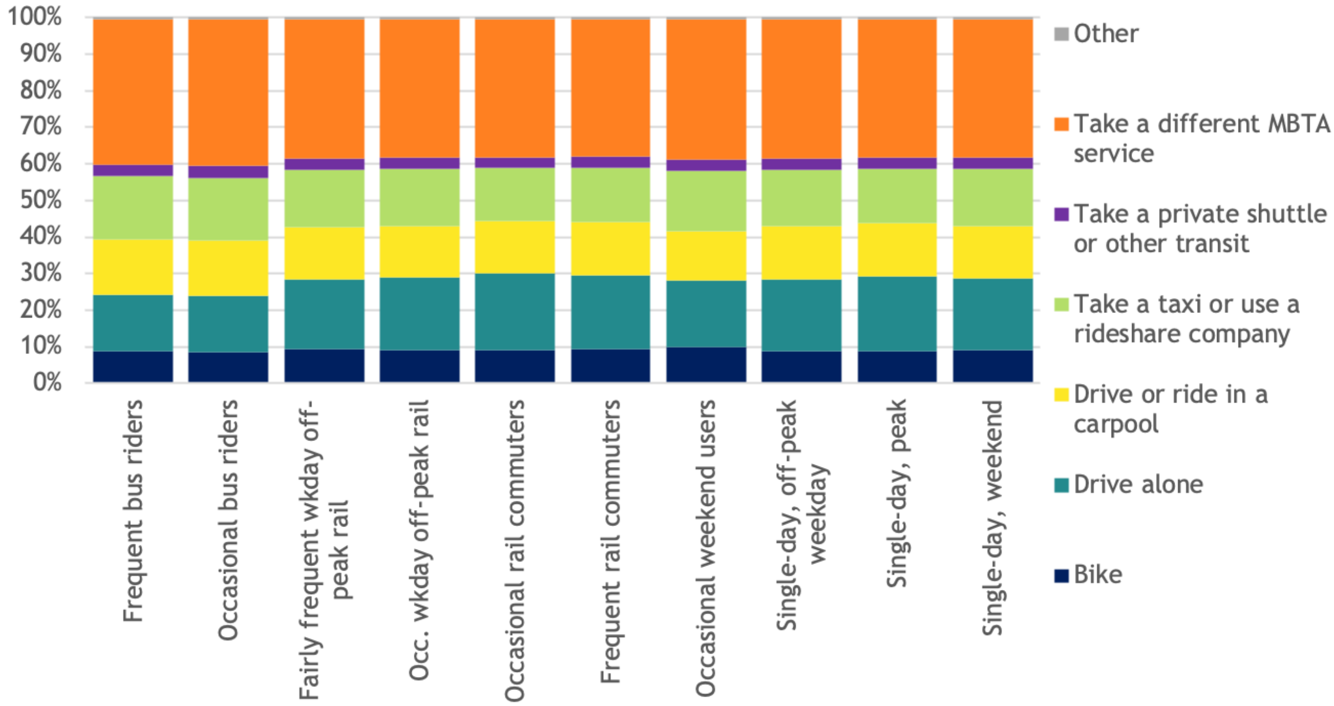


Figure 4-30: Estimated access to alternative modes by cluster based on MBTA survey data

On the other hand, occasional rail commuters are the most likely to have one or more usable household vehicles, with only 27% responding that they had no access to a usable household vehicle.

Finally, Figure 4-32 assesses the distribution of modes used to access the entry point to the journey that the respondent was undertaking at the time of being interviewed for the survey. For all clusters, walking or biking is the dominant first-mile access mode, with at least 86% of any cluster extrapolated to be using these modes. Frequent bus riders, occasional bus riders, and occasional weekend rail riders have the highest usage of walking and biking for first mile access (over 90%).

In summary, the System-wide Passenger Survey results suggest that bus riders are least likely to have car access, a factor that may translate to higher transit dependence even during the pandemic.

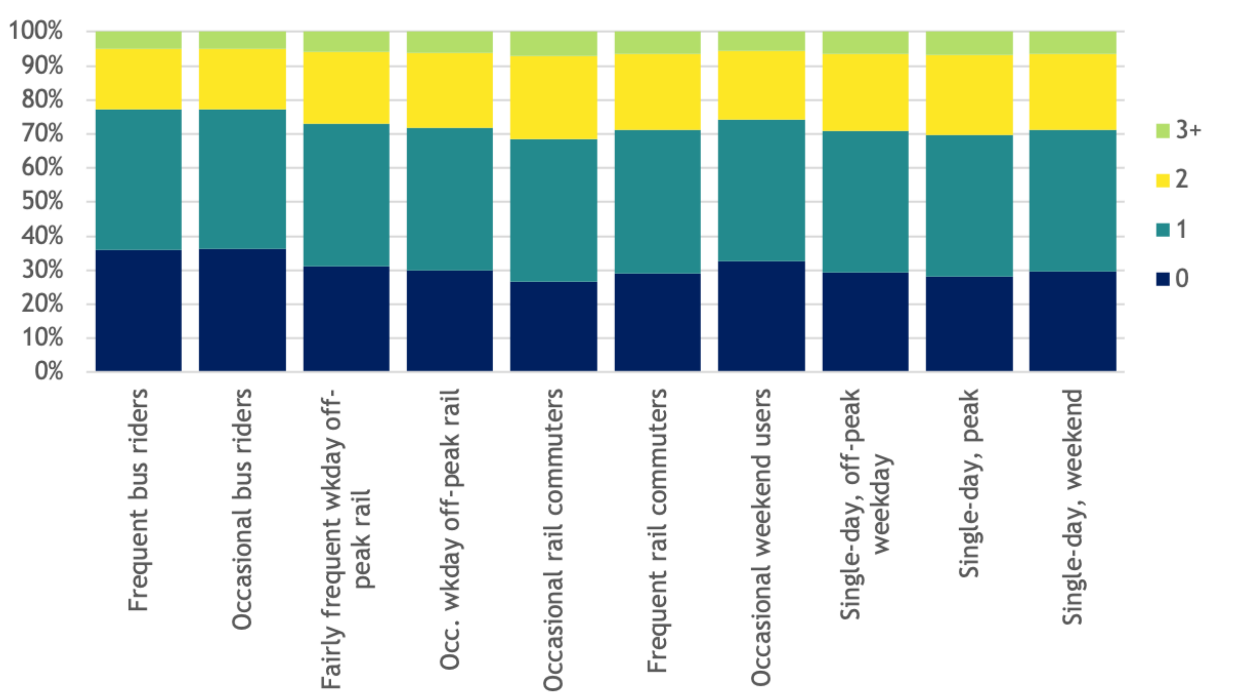


Figure 4-31: Estimated household vehicle count by cluster based on MBTA survey

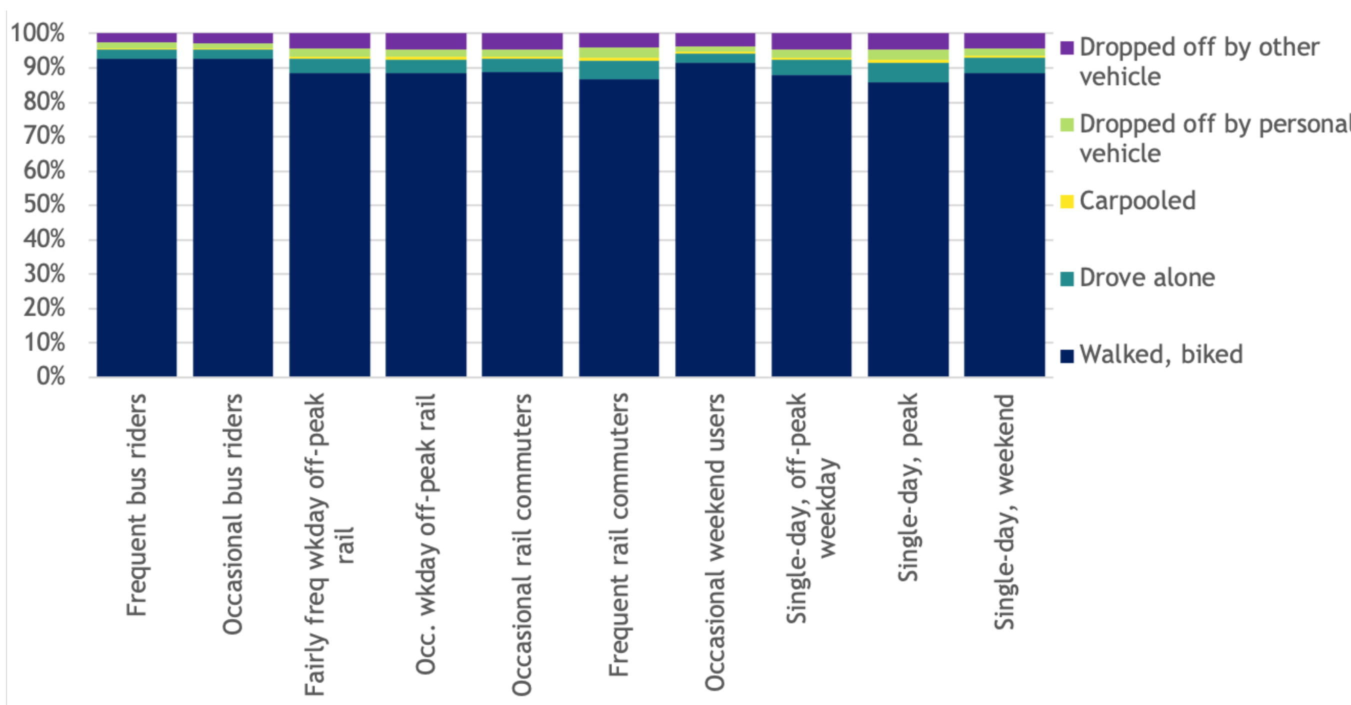


Figure 4-32: Estimated distribution of modes used to access first transit entry point, by cluster based on MBTA survey

4.7 Leveraging baseline clusters for COVID-19 era transit analysis

This chapter paints a picture of pre-pandemic ridership behavior, discussing how distinct rider groups can be in terms of their modal usage, temporal travel patterns, and geographic distribution. Tracking each rider cluster's evolution in COVID-19 will give insight on how the distribution of demand on the MBTA system evolves during this crisis and how the evolving transit demand is distributed across riders from full fare, reduced fare, and corporate-sponsored fare product categories. It will also help target recovery by providing insight on the pattern of transit services reduced-fare populations need, and services that revenue-driving rider clusters need to continue contributing to MBTA coffers.

Chapter 5

COVID-19 impact on ridership behavior

5.1 Overview

Massachusetts was one of the earliest states in the U.S. to report positive COVID-19 cases, with the first case confirmed on February 1, 2020 [69]. Governor Charlie Baker announced the state's first stay-at-home order on March 23 for a duration of two weeks as cases began to skyrocket towards its first peak in spring 2020 (Figure 5-1) [70].[1] During this time, public transit ridership began to plummet, especially for train which fell 94%. The exact size of the collapse in bus ridership is known with less certainty, since the implementation of rear-door boarding from March 21 through July 20 eliminated data flows to fareboxes and thus automated fare collection (AFC) data.[2] Once front-door boarding and fare collection recommenced, Origin-Destination-Transfer (ODX) data derived from AFC indicates that rail trip stage volumes had fallen 86% from our winter baseline (January 13-February 16, 2020) to our COVID analysis period (August 18-September 21, 2020), compared to 63% for bus (Figure 5-2).

[1]The figure is for Suffolk County, which covers Boston but not all areas of Metro Boston, such as Cambridge, MA.

[2]Automatic passenger counting (APC) data was still available, but only for a subset of buses as APC rollout was still in process during this time.

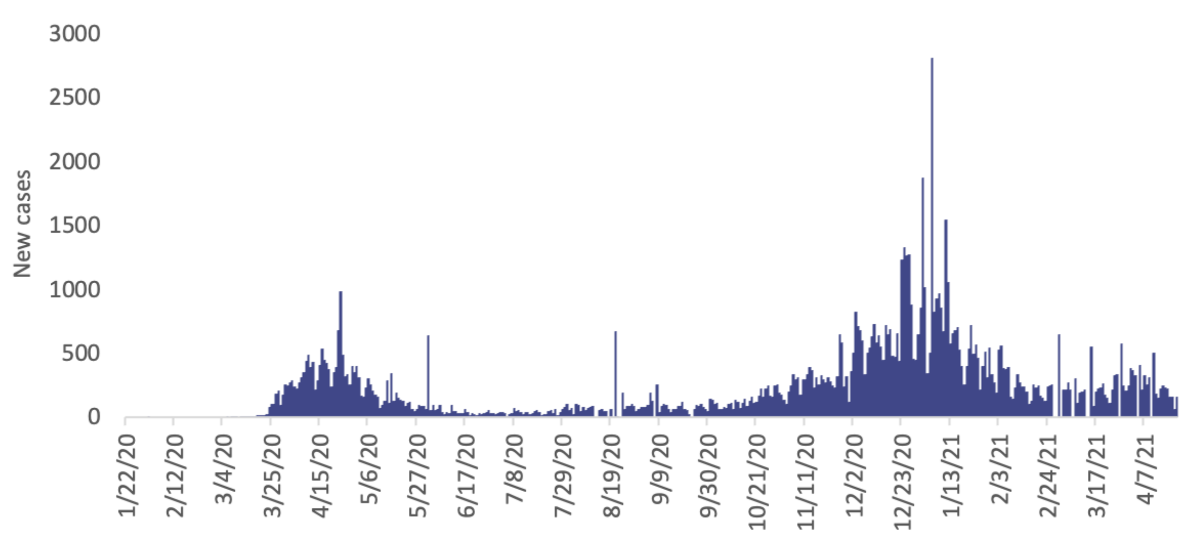


Figure 5-1: Suffolk County COVID-19 cases, January 2020 - April 2021

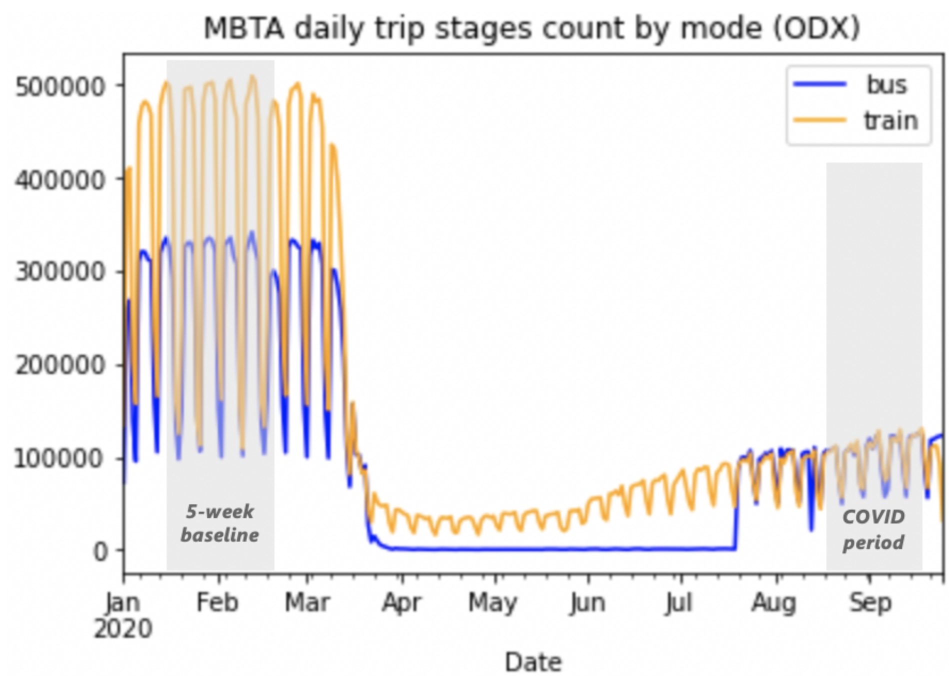


Figure 5-2: MBTA transit ridership trends in 2020 (ODX), study periods shaded

We selected August 18 - September 21, 2020 as our COVID-19 study period because of the data limitations imposed by the rear-door boarding public health policy, and to leave enough time after the end of that policy for front-door boarding and fare payment to fully resume. With five months of COVID-19 experience, there was also greater stabilization of pandemic mobility patterns among travelers in Metro Boston [4]. However, caseloads were at record lows during this time frame, the weather warm, and a portion of students were returning to university campuses, all of which may be associated with the slight uptick in transit usage for both bus and rail. At the same time, the MBTA was reducing the frequency and hours of service on the supply side in response to COVID-19, as detailed in Chapter 6. This thesis takes a demand-side approach, which means it does not explicitly account for the impact of supply change and only captures the changes' indirect effect on demand.

This chapter begins by taking an overview of ODX data during the summer COVID-19 study period, where we find that peakiness in the weekday temporal profile for transit was heavily attenuated but did not disappear with the onset of COVID-19 and remote work. The chapter then presents clustering results from applying the baseline 2020 k-means model from Chapter 4 to COVID-period data, and discusses the shifts in behavioral clusters observed during the pandemic. Subsequently, it turns to the subset of pandemic-period riders who were also in the pre-COVID baseline period—that is, riders who continued riding through the pandemic. This allows us to assess the share of passengers from the baseline who did not churn, and to track how these retained members of each baseline cluster *modified* their behavior during the pandemic such that they shifted over to a separate cluster by the COVID period.

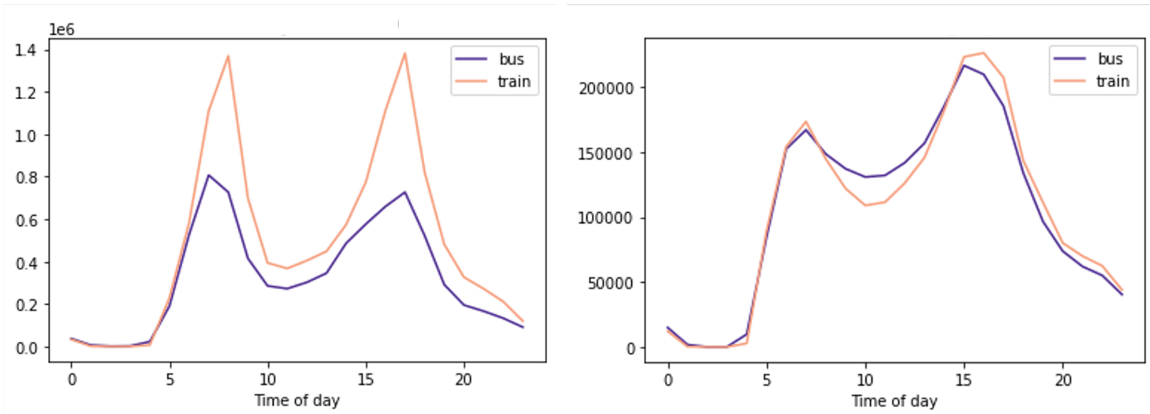
We find that during the pandemic, the share of passengers following the travel profile of frequent bus riders held more steady than could be expected given the "background" cluster-switch rates observed in 2019. This cluster also became a larger contributor to journey count than during even the baseline period. Tracking baseline period frequent bus riders, we find that they were more likely than those from any other baseline cluster to extend their existing ridership behavior into pandemic times. Additionally, fairly frequent weekday off-peak rail riders also made up a slightly larger share of the summer rider and journey count than background rates would suggest. In stark contrast, all clusters representing rail commuters collapsed, likely bringing with it a nosedive in Perq revenues.

Results from this chapter suggest that, though it is unclear how transit ridership overall will recover from COVID-19, greater certainty exists for the travel needs of bus-oriented riders and to a lesser degree fairly frequent off-peak rail riders. The recovery of both frequent and occasional rail commuters, and the significant Perq pass revenue they used to bring for the MBTA, face enormous uncertainty. Addressing this may require the agency to work more directly with institutions subscribing to Perq to understand the timing and design of return-to-office plans, perhaps influence those plans, and reshape the Perq program to keep it cost-effective for Perq workers given their potentially reduced commuting needs.

5.2 COVID period 2020 ODX data: Temporal profiles of transit use

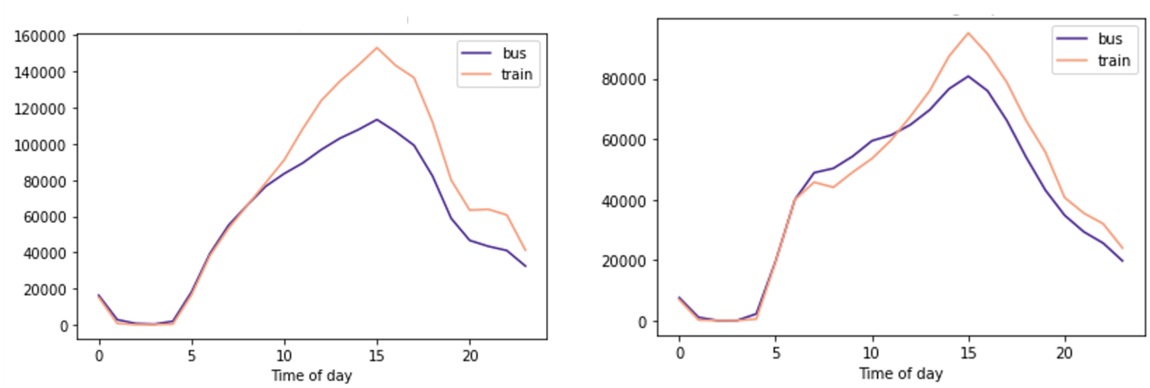
This pandemic analysis relies on ODX data for more direct comparability of results between the pre-COVID baseline and the COVID era. In addition to the cumulative drop in total bus and rail ridership noted in Figure 5-2 above, there were also clear changes in the daily temporal profile of ridership.

Figure 5-3 compares the baseline 2020 temporal profiles of bus and train ridership for a typical weekday, against those from the COVID period. During the pandemic, the peakiness of rail ridership fell precipitously especially for the morning rush hour. In baseline 2020, rail travel was roughly 3.5 time higher in the morning peak at 8 AM or the evening peak at 6 PM than during the midday hour of 11 AM. By the COVID period, this ratio had shrunk to 1.5 for the morning peak compared to midday, and 1.9 for the evening. Notably, train travel volumes during midday were slightly below bus during COVID, whereas it had been above in the baseline. Bus travel also saw reduced peakiness, but the change in the shape of its temporal profile was less drastic



(a) Baseline 2020 (Jan-Feb) (b) COVID-19 2020 (Aug-Sept)

Figure 5-3: Temporal profile of weekday transit use by mode, baseline v COVID-19 (note the difference in the y-axis scale, which is in units of trip stages)

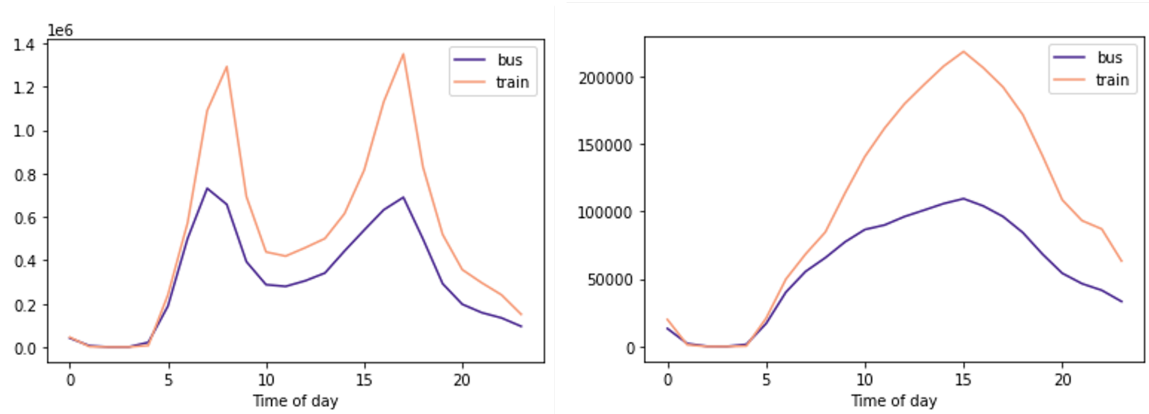


(a) Baseline 2020 (Jan-Feb) (b) COVID-19 2020 (Aug-Sept)

Figure 5-4: Temporal profile of weekend transit use by mode, baseline v COVID-19 (note the difference in the y-axis scale, which is in units of trip stages) since bus travel already had significant off-peak, midday ridership during the baseline. Similar to rail, bus ridership saw greater shrinkage for the morning rush hour than for the evening, possibly because of the continuation of some evening social activities even as commutes dwindled with the transition to remote work. The ratio of morning peak bus ridership at its high point at 7 AM to the midday nadir at 11 AM was 2.7 during the baseline, and the ratio for the evening peak at 7 PM was 2.6. Once the pandemic hit, these ratios shrank to 1.2 and 1.6 respectively. Unlike in some transit systems such as Chicago's, the peaks did not disappear completely, especially for the evening rush hour [32].

The weekend temporal profile of travel underwent much less change during COVID-19, especially among bus riders (Figure 5-4). Volumes shrank by roughly half instead of the close to 80% seen during the weekday. Bus travel maintained a similar temporal distribution throughout a typical weekend day, though the rise in journey starts between 5 AM and 8 AM is steeper than during the baseline. There is also a clearer evening peak. On the other hand for trains, midday seems to capture a lower share of trip stages than during the baseline, leaving the evening peak sharper by comparison. Yet the evening peakiness of weekend train travel is no longer as different from that of bus riders. These temporal profiles already hint at the differential impact of COVID-19 on rail versus bus ridership, and the operational and scheduling impacts of shifting behaviors especially when it comes to planning for weekends when frequencies tend to be lower and schedules more evenly spaced.

In addition to comparing the COVID 2020 temporal profiles against the pre-COVID baseline, we also check against a similar time period of summer 2019 (Figure 5-5). Since we noted seasonal differences in ridership in Section 3.5, it is helpful to check COVID 2020 data against a previous period in the same season in order to control for the "background" differences between winter and summer ridership patterns. Figure 5-5 indicates that a typical summer's temporal profile for weekday or weekend ridership much more closely resembles the baseline 2020 profile than the pandemic-period summer 2020 profile, with an even clearer gap between weekend evening train versus bus ridership. The COVID 2020 shifts in how travel is distributed throughout



(a) Weekday summer 2019 (b) Weekend summer 2019

Figure 5-5: "Background" temporal profile of summer transit use by mode, summer 2019 (note the difference in the y-axis scale, which is in units of trip stages)

the day therefore appear to reflect a pandemic situation rather than natural seasonal variation.

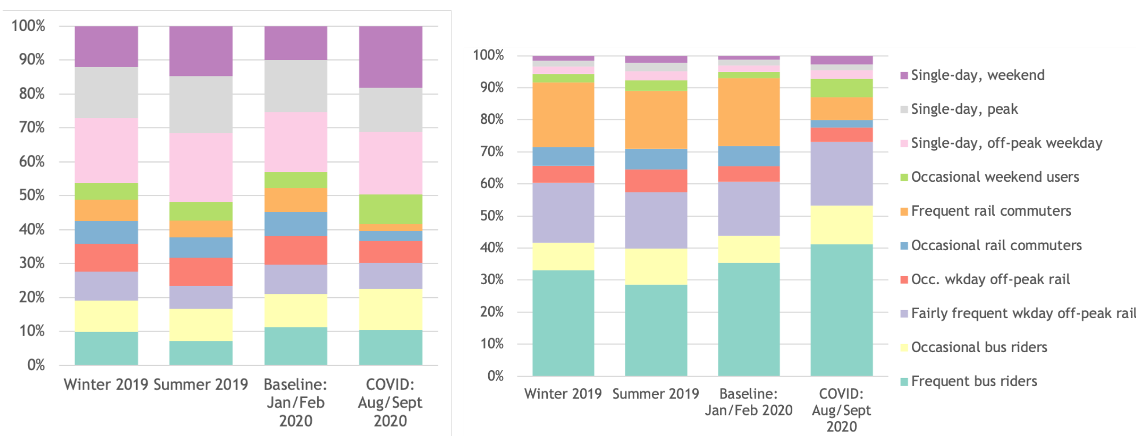
In Chapter 4, cluster profiling indicated that the time of day of ridership was a key distinguishing feature among rail riders and also among single-day riders; there were notable modal differences in ridership patterns as well. These factors also appear prominent in the aggregated temporal and modal differences observed in how ridership changed during the pandemic. In the next two sections, we drill down to the cluster level and assess how clusters of passengers that differed in their pre-COVID transit use patterns may have diverged in adapting their mobility behavior to the pandemic.

5.3 Clustering passengers based on pandemic transit usage behavior

In this section, we examine the pandemic's impact using the behavioral clustering model trained on winter 2020 data, to assess 1) changes in the composition of behavioral clusters present among pandemic riders, and 2) changes in the distribution of feature values within each cluster from the baseline to the COVID-19 period. For this portion of the analysis, we take *all* passengers using the MBTA system during the COVID-19 period in 2020, and assign them to behavioral clusters based on their continuing travel behavior. This is an out-of-sample application of the clustering model trained on baseline 2020 data, since not all riders in the COVID-19 period were present in the baseline.

5.3.1 Passenger and journey composition

Figure 5-6a and Figure 5-7 show how the allocation of passengers across the ten behavioral clusters evolved from baseline 2020 to the COVID-19 period in summer 2020, and compare that against the "natural" seasonal change from winter to summer 2019. Figure 5-7 (and the corresponding Figure 5-8 for journeys), explicitly focuses on this "difference-in-difference" for each cluster—that is, we examine the difference in bar heights between the blue 2019 data, and compare that to the difference in bar heights between the warm-colored 2020 data.



(a) Passenger composition (b) Journey composition

Figure 5-6: Cluster composition for winter 2019, summer 2019, baseline 2020, COVID-period 2020

In the COVID-19 period, the three clusters most associated with reduced fare users—that is, the frequent bus riders, occasional bus riders, and fairly frequent weekday off-peak rail riders—provided roughly the same share of the passenger pool as in baseline winter 2020. This is in stark contrast to the 2019 seasonal pattern where these groups shrank as shares of the passenger count after winter.

In the pandemic, 10.4% of passengers were frequent bus riders, 12.1% occasional bus riders, and 7.8% fairly frequent weekday off-peak rail riders. This reflects a strong

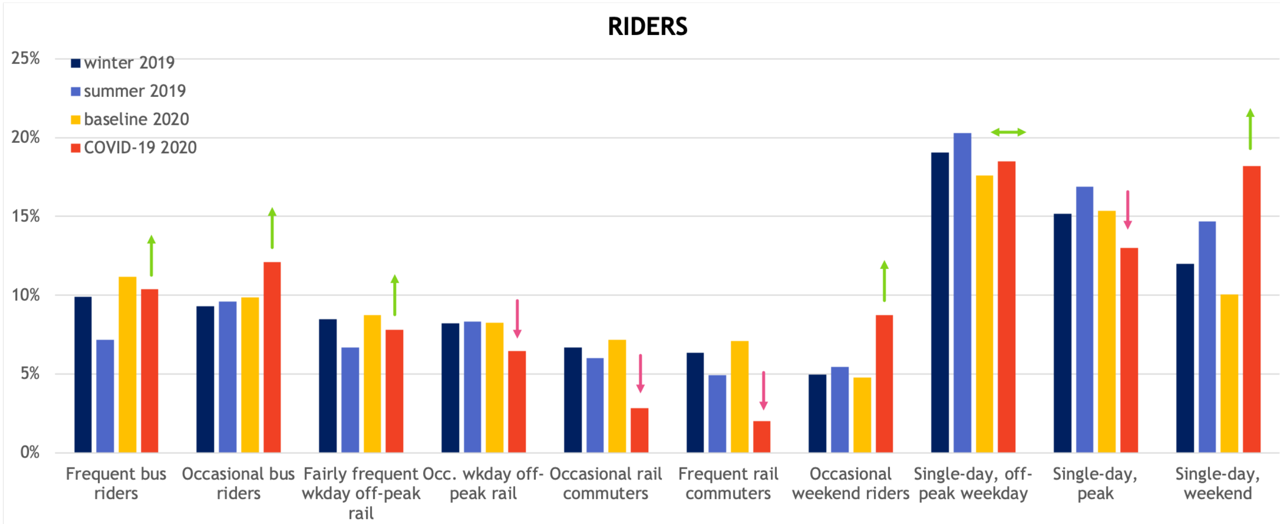


Figure 5-7: Passenger composition change, baseline/COVID 2020 compared to corresponding seasons of 2019

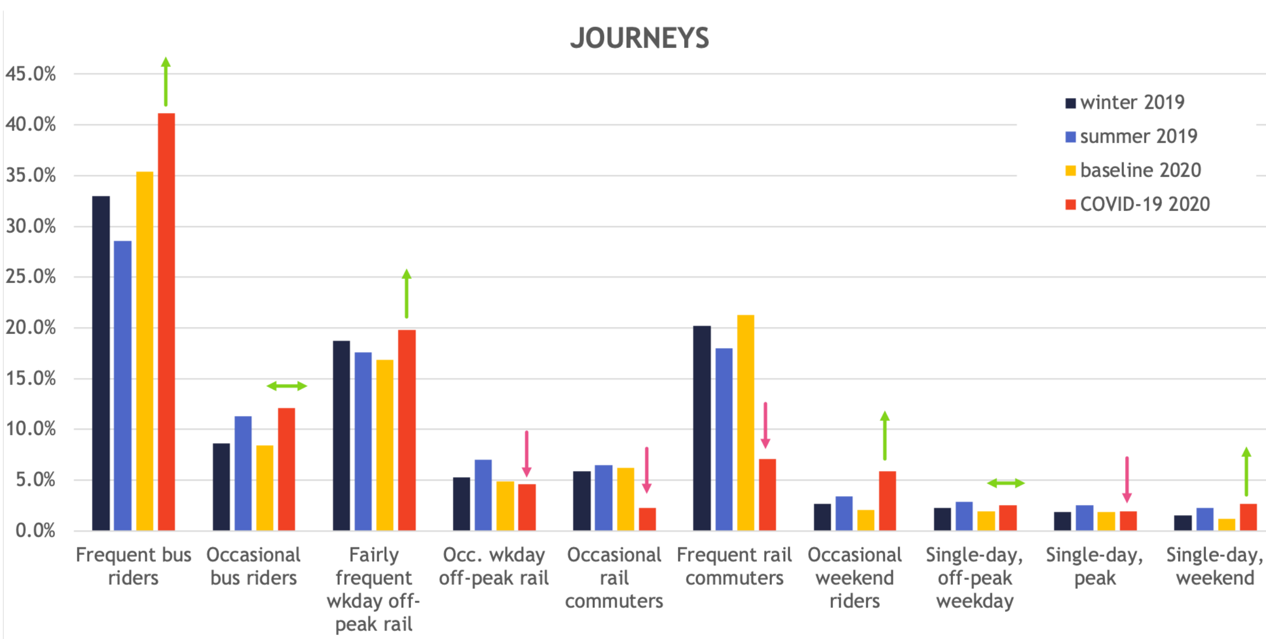


Figure 5-8: Journey composition change, baseline/COVID 2020 compared to corresponding seasons of 2019

shift towards occasional bus ridership after the baseline, one that appears larger than can typically be expected given summer 2019, when the passenger share of this cluster (9.6%) was close to that of both winter of the same year (9.3%) and baseline 2020 (9.9%). Further, the share of frequent bus riders (10.4%) shrank only 6.9% from the baseline (11.2%), in contrast to 2019 when frequent bus riders' share shrank 27% from 9.9% down to 7.2%. As a result, the frequent bus riders' share of passengers is 44.5% larger than in the same season of 2019.

Fairly frequent weekday off-peak rail riders (7.8%) were also slightly down from the baseline (8.7%) but larger than in the corresponding season of 2019 (6.7%) where there was a stronger fall-off from winter to summer. The occasional weekday off-peak rail riders shrank to 6.5% during the pandemic, down from 8.2-8.3% observed in the other three time frames. The single-day off-peak weekday riders, who are mostly rail users, rose to 18.5% during the pandemic from 17.6% in the baseline, but this seems close to being in line with the winter-summer evolution observed in 2019.

As a segment, the single-day clusters surged between the baseline and summer. However, the composition of these clusters tilted in favor of weekend single-day riders (18.2%), much more so than was expected based on 2019 when the cluster expanded from 12.0% to 14.7%. This aligns with the notable rise in multi-day weekend rail riders, which expanded from 4.8% in the baseline to 8.7% in the pandemic summer, outstripping the 2019 situation where this cluster held relatively flat inter-seasonally.

At the same time, the single-day peak period cluster declined from 15.3% of baseline period riders to 13.0% in the pandemic period, which is in contrast to the uptick observed from winter to summer 2019. This aligns with the virtual disappearance of the two commuter clusters during the pandemic. Frequent rail commuters collapsed to 2.0% of all passengers compared to 7.1% in the baseline. In 2019, the winter to summer change was much more gradual, from 6.3% down to 5.0%. Occasional rail commuters shrank to 2.8% during COVID, from 7.1% in the baseline. Again, this is starkly sharper than the inter-seasonal change experienced in 2019, when this cluster's passenger share only dwindled slightly from 6.7% to 6.0%.

In summary, the two bus clusters and the two weekend travel clusters most clearly contributed a greater share to total passenger count during the pandemic period than could be expected if we assume that 2019's inter-seasonal ridership patterns held. Commuters who use rail vanished almost entirely, while weekday off-peak travel remained somewhat steady.

Figure 5-6b and Figure 5-8 complement the above passenger perspective by examining the evolution of each clusters' contribution to *journey* count between the time windows under consideration. As noted in Chapter 3, journey counts are more essential to operational planning than passenger count, as it represents the actual demand put upon the transit system.

As we move from summer 2019 to baseline 2020 and finally summer 2020, there is a consistent rise in the share of each period's journeys made by frequent bus riders. Whereas this cluster contributed 35.4% in the baseline, 33.0% in winter 2019, and only 29.0% in summer 2019, it was 41.2% of all journeys in the pandemic period. Thus in contrast to the inter-seasonal pattern in 2019, frequent bus riders took a larger share of the total journeys in the pandemic than in the baseline. This is perhaps especially striking given that the share of pandemic-period passengers categorized as frequent bus riders was slightly lower than the baseline period's, not higher. If we look instead at the absolute value of the journey and passenger numbers, it appears that this is occurring because the average number of journeys per frequent bus rider stayed flat during the pandemic, while this ratio fell for some of the other large clusters.

Among occasional bus riders, an uptick in the share of total journeys was observed as well from 8.4% in the baseline to 12.1% during the pandemic. However, this is not a notable departure from the inter-seasonal pattern seen in 2019. Given the uptick in the share of *passengers* categorized as occasional bus riders during this time, this reflects that those in the growing pool of occasional bus riders remained mostly in line with historical expectations or slightly reduced their journey count.

From the rail-oriented clusters, the fairly frequent weekday off-peak rail cluster saw its share of total journeys rise from 16.9% to 19.8% in the pandemic, in contrast to 2019 when its share shrank from winter to summer. Similar to the frequent bus riders cluster, this increase in the relative share of journeys from baseline to summer 2020 comes alongside a fall in the share of riders, suggesting that journeys per passengers is also rising for this group. The other two off-peak clusters with less intensive usage did not, however, show this surge in their contributions to total journey count.

In contrast with the frequent bus riders and fairly frequent weekday off-peak riders, both rail commuter groups again saw a precipitous fall in their contributions to journey count once the pandemic hit. In 2019, frequent rail commuters were 20.2% of journeys in the winter and slightly less at 18.0% in the summer. By contrast in 2020, this cluster began at a higher 21.2% in the winter baseline but finished at one-third that level, 7.1%, once the pandemic hit. Occasional rail commuters actually saw an increase in its journey share in 2019 going from winter to summer, but this collapsed by two-thirds across the corresponding periods of 2020.

Finally, enforcing the previously observed trend in the relatively stronger performance of weekend travel during the pandemic, the occasional weekend riders experienced a tripling of its share of total journeys from the baseline into the pandemic summer, far above 2019 when it only rose from 2.7% to 3.4%. Single-day weekend riders also saw a slight uptick compared to 2019.

Overall, journey-level results suggest that passengers who behave like frequent bus riders during the pandemic were the notable cohort actively riding transit during this emergency period, and that per passenger, their journey volumes may have risen slightly. Those behaving as occasional bus riders were also more likely than most other clusters to still use transit during COVID-19, but their average usage intensity fell. Meanwhile, demand from rail commuter clusters was pummeled by the pandemic. Lastly, passengers using transit for the weekend—likely, for leisure purposes—became a larger share of the overall smaller rider and journey pool, but are still insignificant contributors to journey volumes.

These observations already suggest the MBTA may need to continue serving the temporal and spatial mobility patterns of frequent bus riders, who may be transit dependent or essential workers given their high usage of transit during pandemic times. Mid-day rail travel is another category where service needs to be maintained. Meanwhile, rail usage, especially for commuting, has transformed entirely due to remote work, and depending on the evolution of the workweek in the COVID-19 recovery it may not return to the same state as winter 2020. Therefore, there is a higher level of certainty in how the most intensively used bus services and midday rail services will look going forward, than there is certainty on rail commuting.

5.3.2 Feature distribution by cluster, baseline versus COVID-19

Throughout this chapter's analysis, we apply the k-means model trained on baseline 2020 ODX data to tease out distinct behavioral patterns in the COVID-19 era data. This means the centroids of each pandemic period cluster we are assessing is the same as during the baseline. However, the distribution in feature values among the points categorized into each cluster may still differ from the baseline—i.e., the distribution of points in each cluster around each centroid can shift, even if the centroids do not.

Figure 5-9 presents violin plots for the features used to cluster multi-day transit riders during the COVID-19 period in 2020. It is analogous to the baseline period plot in Figure 4-7 of the previous chapter. Comparing these two figures reveals an upward shift in the bus share of trip stages among the frequent and occasional bus riders still using the MBTA system; the upward shift is evident in both the IQR and the median value (clusters #1, #2). This is accompanied, however, by a slight downward shift in transfers among these two groups.

Temporally, there is a decrease in peak-period travel among frequent bus riders (cluster #1) as well as both frequent and occasional rail commuters (clusters #6, #5). At the same time, the active range and active days distributions are shifting upward for frequent bus riders. Frequent rail commuters shrank drastically as a group, but those still riding following this commuting pattern are still highly engaged with the system in terms of both active range and active days. Weekend travel shares were also notably higher for this rider cluster, as well as for fairly frequent weekday off-peak rail riders, frequent bus riders, and occasional bus riders. In fact, whereas the weekend usage distribution in the baseline had offered a clean-cut contrast between

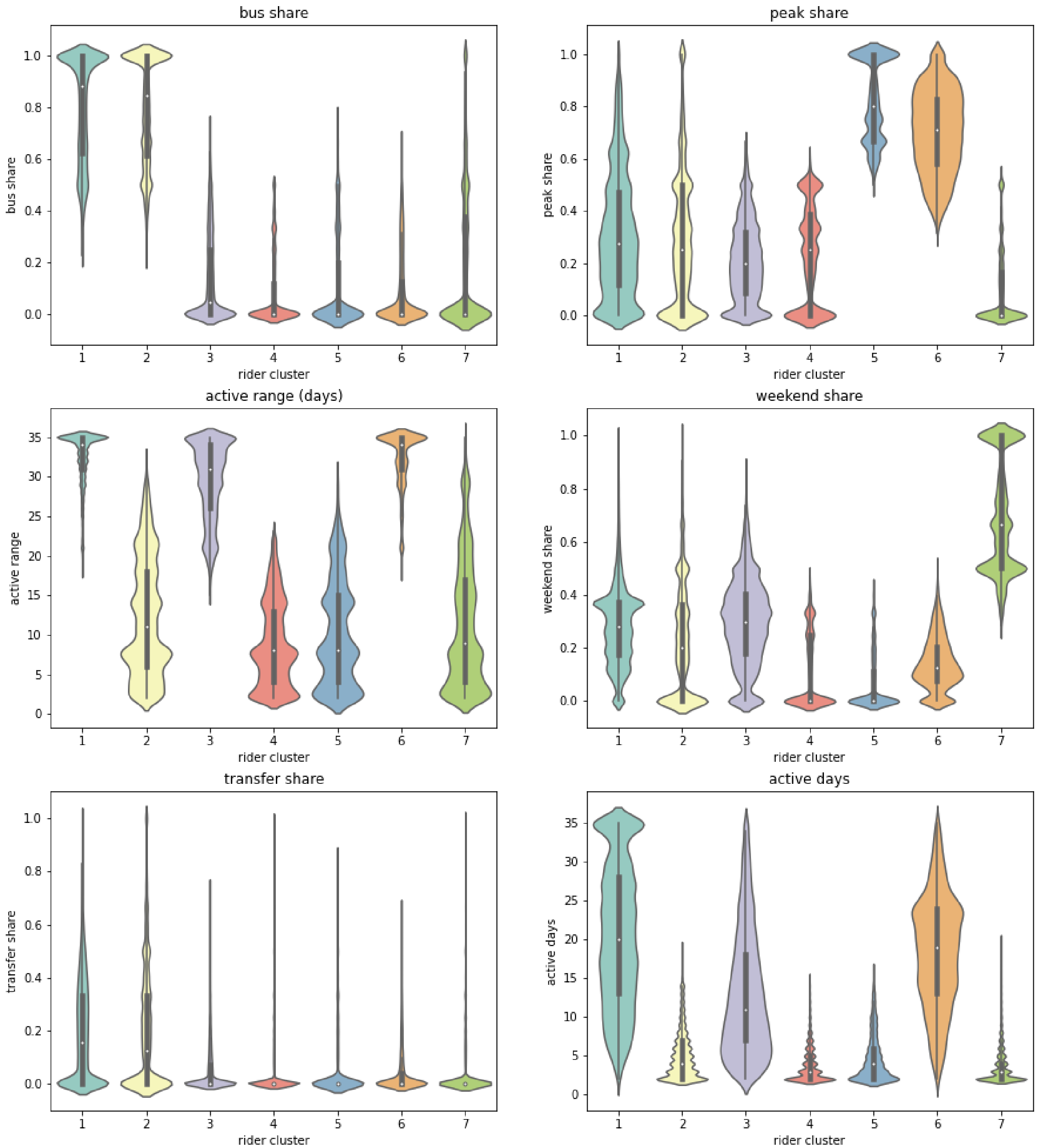
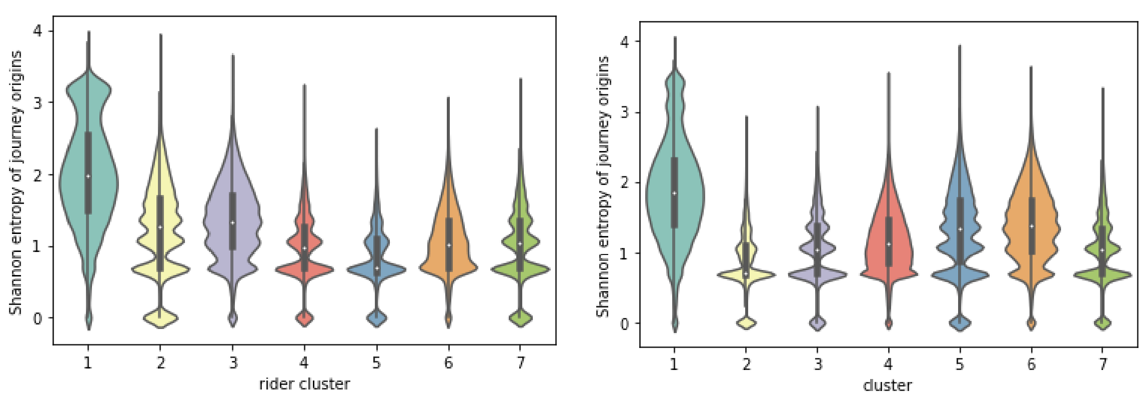


Figure 5-9: Variation of features within and between multi-day rider clusters during COVID-19

the weekend rail riders cluster (#7) and the other multi-day groups, during COVID-19 the distinction is less clear cut due to the upward shift along this dimension in nearly all multi-day clusters.

Figure 5-10 compares the entropy in boarding locations of passenger clusters during COVID-19 against the baseline. Among frequent bus riders the distribution of entropy scores among members shifted upward compared to previously, as average journeys per passenger also ticked up. This suggests that those still using the bus intensively during COVID-19 are boarding more evenly across a range of locations (rather than strongly favoring just a handful of stops). Examining top boarding stops/stations for bus riders also suggests relatively high stability between the baseline and the COVID-19 period. Among occasional bus riders and fairly frequent weekday off-peak rail riders, the latter of which saw a slight uptick in average journeys per passenger during COVID-19, there are also increases in entropy. Meanwhile, the entropy distribution shifted down for the frequent rail commuter cluster as well as the remaining rail groups. These observations mostly enforce our earlier discussion on the operational need to pay close attention to the existing mobility demand of frequent bus riders, and also of rail users intensively traveling off-peak.



(a) COVID-19 period summer 2020 (b) Baseline period winter 2020

Figure 5-10: Entropy of journey origins among multi-day rider clusters

5.3.3 Cluster profiling

The previous chapter discussed the revenue importance of Perq riders. Given the collapse of frequent rail commuters during COVID-19, this revenue stream is under threat. Riders on Perq passes shrank 73.7% from the baseline to summer 2020, faster than the 60.4% overall shrinkage in passenger count. Figure 5-11, which plots the share of Perq riders in each cluster, shows that in particular, frequent rail commuters is no longer dominant in its Perq passenger contributions. This group used to provide 39% of Perq passengers in the baseline, far exceeding frequent bus riders in second place with 29% and fairly frequent weekday off-peak rail riders at 15%. It now only contributes 17% of Perq passengers, behind both frequent bus riders (27%) and fairly frequent off-peak riders (23%). Frequent rail commuters' Perq passenger contributions are still, however, punching far above its weight given how few riders of that category are left in the system (2%), suggesting that there is a sub-category of institutionally sponsored rail riders who remain consistently engaged with transit.

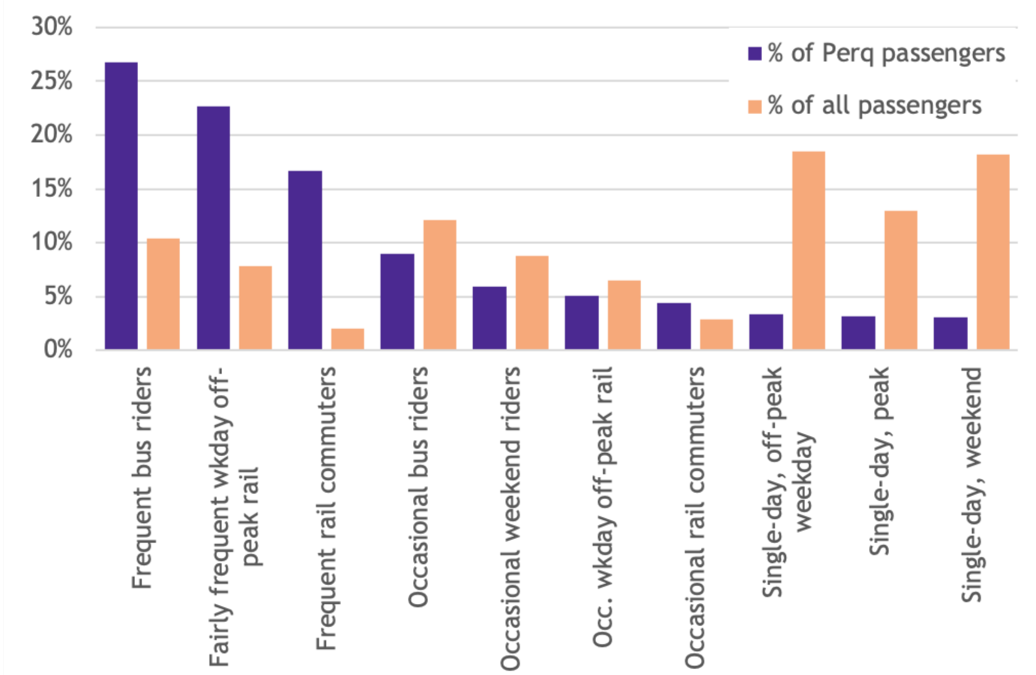


Figure 5-11: Cluster contributions to Perq v. general passenger pool, COVID-19 summer 2020

Other smart card meta-data reveal a general drop-off in the share of each cluster using monthly passes, accompanied by a shift towards pay-as-you-go. Frequent rail commuters and fairly frequent weekday off-peak rail riders were the two groups where the share of monthly passes remained relatively comparable to baseline (falling 62% to 50% and 30% to 27% respectively).

As Metro Boston recovers from COVID-19 and large employers reassess the extent to which remote work becomes permanent, there may be further changes on the horizon for Perq rail commuters and the revenue this group can promise for MBTA operations. In the meantime, we observe again in the Perq setting that the degree of uncertainty in ridership is lower among frequent bus riders and fairly frequent off-peak rail riders, suggesting that even during the dynamic times of the pandemic and subsequent recovery, transit planning efforts targeting these groups may be able to continue without as drastic an overhaul as efforts targeting rail-based commuter groups. There also appears to be a small sub-group of frequent rail commuters who continue to exhibit high and consistent usage during emergencies, but they are currently a much smaller share of passengers than that covered by the other two clusters.

Smart card meta-data regarding user type shows that during the pandemic, the frequent bus rider, occasional bus rider, and fairly frequent weekday off-peak rail rider continued to be the top behavioral clusters for those on reduced fares (Figure 5-12). However, the composition of reduced fare users remaining in the system changed. The share of TAP user types rose in these three clusters. Seniors rose as a share of frequent bus riders (from 6.4% in the baseline to 9% during COVID-19)—and also as a share of frequent rail commuters (2.6% to 4%).

Students vanished from the frequent bus rider, frequent rail commuter, and fairly frequent weekday off-peak clusters, while also falling as a share of occasional bus riders from 6% to 4%. Meanwhile, the remaining students shifted towards the occasional off-peak rail and occasional rail commuter clusters, showing a lessening intensity of previously observed transit usage patterns. Further, there was a doubling of the Other reduced fare category among remaining riders who are frequent bus users; this was driven almost entirely by short fares. Adult fares, on the other hand, held relatively steady as a percentage of each behavioral cluster.

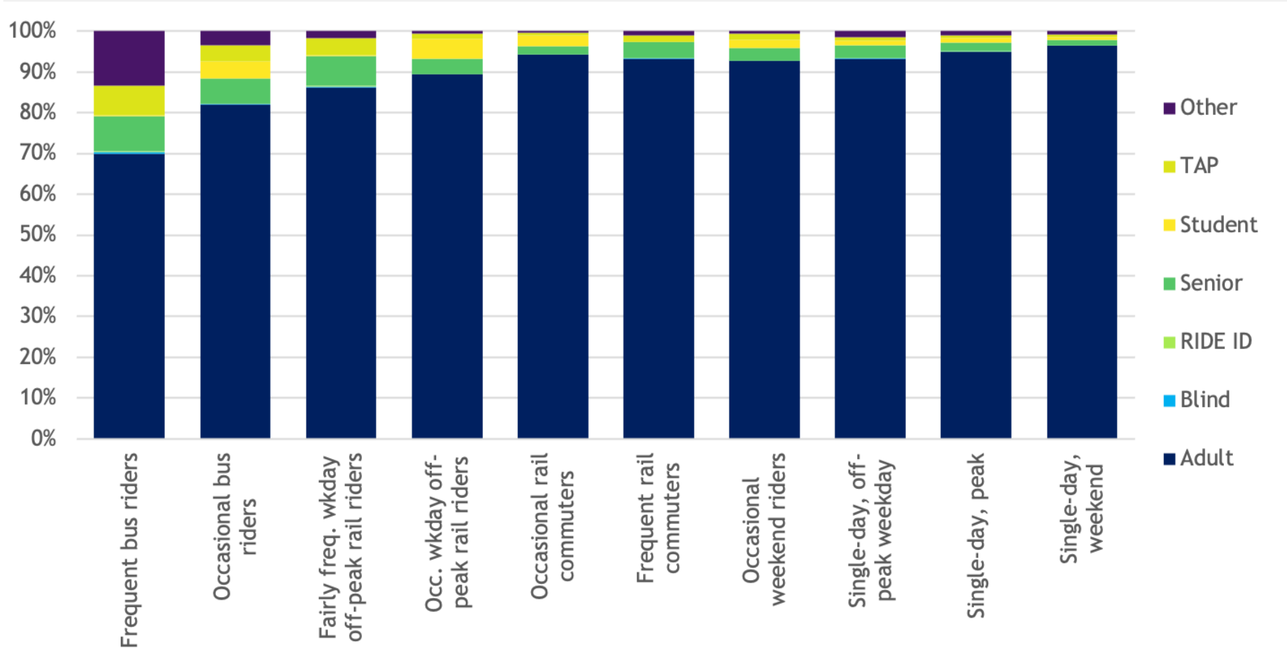


Figure 5-12: Clusters broken down by user type, COVID-19 summer 2020

Figure 5-13 flips the previous analysis to assess the contribution of clusters to each reduced fare type. This perspective highlights the shift in adult full fare-paying riders away from the frequent bus rider, frequent rail commuter, occasional rail commuter, and single-day peak rider clusters towards the occasional bus rider, occasional weekend rider, and single-day weekend clusters. This again emphasizes the overarching movement away from regular work-based, consistent transit use towards more occasional and leisure-oriented travel during the pandemic.

The occasional weekend cluster rose in importance not only for adult full fare paying riders, but for students as well. Figure 5-13 makes the change in student ridership behavior even clearer, showing a movement almost completely away from frequent bus ridership (down from 44% to 1% from baseline to the COVID period) towards occasional bus ridership (up from 15% to 32%). Students also no longer rode frequently in the weekday off-peak times, shifting instead to occasional off-peak use (up from 3% to 20%). Occasional weekend (up from 4% to 12%) and the three single-day ridership clusters also saw dramatic expansions (up from 9% to 31%).

Finally, a larger share of seniors continuing to use transit during the pandemic are shifting towards bus, expanding both the frequent bus rider cluster (up from 19% to

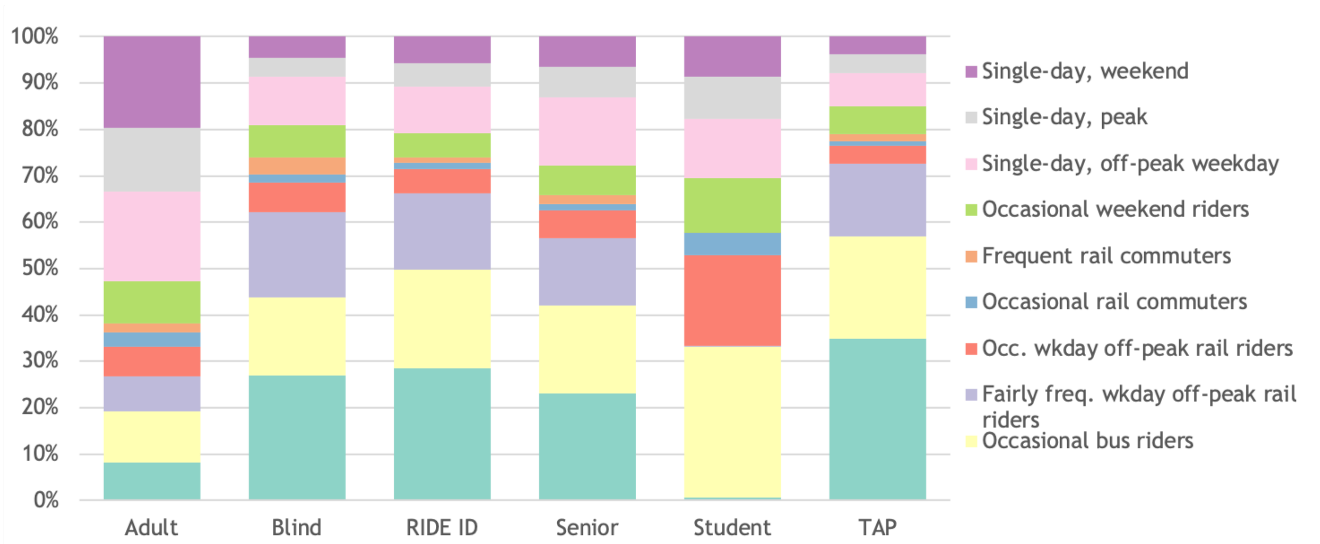


Figure 5-13: User type broken down by clusters, COVID-19 summer 2020

23%) and the occasional bus rider cluster (15% to 19%). Among other major reduced fare categories, the distribution of ridership patterns are more stable.

We do not join our summer ODX data to the 2015-2017 System-wide Survey as we did in the baseline to gather demographic information. The survey's sampling methodology produced statistics that roughly reflected the demographics of those traveling in non-pandemic times. Unlike ODX and associated smart card meta-data, this survey effort does not continue into the pandemic era to capture shifts in ridership during COVID-19. However, taking the high-level results we found from the survey data in the previous chapter, it appears there may be associations between the bus rider clusters' lower household vehicle access and lower income status (which is tied to essential work that cannot be done remotely) that can potentially drive the continuation of bus ridership observed during the pandemic. Fairly frequent weekday off-peak riders, though they also contain a higher share using reduced fares, appear to have greater access to cars and alternative travel modes, and during the pandemic we see a more muted volume change from this cluster.

In summary, cluster profiling using ODX smart card data showed that Perq, the largest single source of fare revenues, was retaining more traction during the pandemic among frequent bus riders and fairly frequent weekday off-peak rail riders than among its typical bulwark, frequent rail commuters. This suggests that to drive both ridership and fare revenue recovery, the MBTA may need to work closely on the timing of service restoration and the structure as well as pricing of Perq to offer a corporate pass program that can work well with increased remote work. Meanwhile, reduced fare riders continued to make up a larger share of the two bus clusters and fairly frequent weekday off-peak rail cluster than any others, though the composition shifted with students diminishing while senior users rose, suggesting greater continued transit dependence among these user types.

5.4 Churn: Tracking baseline transit passengers through the pandemic

This section takes a more granular look at how transit behavior shifted among those who were already riders in the baseline period, once COVID-19 struck. This differs from the previous section's examination of the COVID-19 summer 2020 riders in that it excludes new passengers who entered the system during the pandemic.[3] Whereas the previous section assessed the overall characteristics of each behavioral cluster during COVID-19, it could not tell us how passenger behavior *evolved* from the baseline to the pandemic period, since it did not track riders longitudinally. Figure 5-14, by contrast, tracks each baseline passenger longitudinally into the COVID period and compares the patterns observed against the background churn and cluster-switching seen in a typical year like 2019.

The left-hand side of Figure 5-14 shows the background rates of churn and cluster-switch during our non-pandemic reference year, 2019. By contrast, the right-hand side of the same figure shows churn and cluster-switch rates from baseline 2020 to COVID 2020. Within each graph, each bar represents one cluster from the winter season, and each colored segment within the bar indicates the cluster that certain members

[3]Technically, we are not excluding new riders, only new smart card numbers not previously see in the baseline. If the same individual rode in baseline 2020, then began riding again in August using a new CharlieCard, such a person would be considered a "new" rider to ODX and AFC.

switched to by the following summer. That is, each bar tracks how passengers of a particular cluster shifted their transit usage patterns through the year. The portion with the same color as the cluster label on the x-axis represents passengers who maintained the same behavioral pattern throughout, and the grey portions represent those who churned (i.e., left the system or were otherwise no longer traceable under the same smart card ID). At least 60% of each cluster churned, with nearly 100% churn for the single-day users.[4]

Frequent rail commuters, fairly frequent weekday off-peak rail riders, and frequent bus riders were the three groups with the highest rider retention (i.e., lowest churn) from winter to summer 2019.[5] During the pandemic, they saw the largest rises in churn, partially because there was more room for retention to fall. The churn rates among these were 51 ppt, 33 ppt, and 24 ppt respectively above the background rates seen in 2019; this also translates to churn rates that were 2.7 times, 1.8 times, and 1.7 times greater than the background rate.

[4]Because of the nearly complete churn among single-day clusters, Figure 5-14 does not include bars for these (i.e., we do not track riders who were in a single-day cluster in the winter). However, the figure keeps single-day clusters as one of the possible "destinations" for where a rider evolved by the summer season, because a notable portion of multi-day riders became single-day especially during the pandemic.

[5]Note that these graphs do not use the same cluster coloring scheme as preceding graphs. Darker colors are used here to enable legibility of the x-axis.

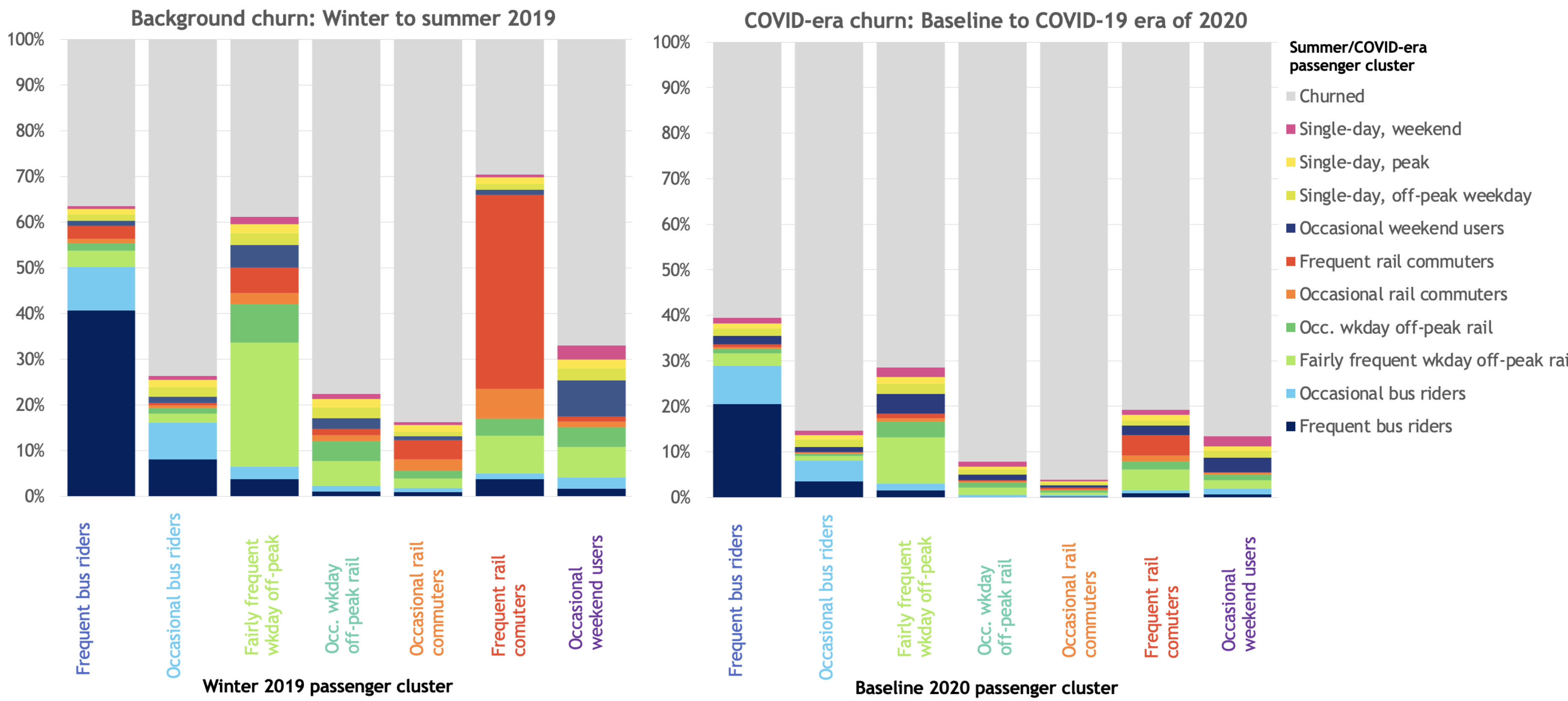


Figure 5-14: Churn and cluster-switch rates, "background" 2019 v. pandemic-period

The drop-off in frequent rail commuters is particularly striking in Figure 3.8, because this cluster used to have the highest passenger retention rate from winter to summer (70%), with 43% even staying within the same cluster in the summer. The frequent bus riders and fairly frequent weekday off-peak rail clusters had slightly lower natural retention rates (63%, 61%). In addition to having the highest uptick in churn, the frequent rail rider cluster experienced the largest decrease in the share of passengers who stayed in the same cluster from baseline to the COVID period (-38 ppt from 43% down to 4.4%), when comparing 2020 figures against 2019 ones. There was even a decrease in the share who became occasional rail commuters (-5.2 ppt) or fairly frequent weekday off-peak rail riders (-3.8 ppt), suggesting that reduction in travel intensity or shifting travel to a less crowded period of the day were not clearly favored strategies. There was also only a small increase (0.9 ppt) in the share who switched to occasional weekend users. Instead, these frequent rail commuters preferred to leave the system; whether because they had eliminated their commutes due to remote work or because they had switched to another mode like driving or biking, we cannot discover from our existing data sources.

Frequent bus riders were the most likely to remain with the same cluster during the pandemic (20.5% stayed in-cluster). Fairly frequent weekday off-peak rail riders were a distance second (10.2%). No other clusters had more than 4.5% of baseline period passengers remain in-cluster during the pandemic. In addition to scoring the highest along this metric, the frequent bus rider cluster is notable in that the share switching to occasional bus riders (8.5% of baseline passengers) was very close to the background switch rate (9.5%) despite the cluster's higher churn during 2020. This seems again to indicate that pre-pandemic passengers who exhibited heavy bus use are the most likely to continue riding transit, and retain more of their typical inter-seasonal behavioral patterns.

Fairly frequent off-peak rail riders, as previously mentioned, were the third most likely to stay in-cluster and retain their typical transit usage patterns during the pandemic. This may be because off-peak rail was already less crowded compared to rush hour travel, diminishing fear of COVID-19 contagion on the enclosed subway cars. The share that shifted to more occasional off-peak ridership is still sizable (3.4%), but is half the background rate; the share that switched to single-day off-peak ridership stayed relatively flat (rising 0.6 ppt). Thus it appears that this cluster of midday rail passengers tended to stay off-peak travelers during the pandemic if they were retained, but sometimes rode with lower frequency.

The occasional bus rider cluster, like all occasional clusters, was already experiencing high background churn. The additional churn in 2020 was therefore less dramatic, at only 1.14 times that observed in summer 2019. Among those in the group who remained, about 24% still increased their transit usage intensity to become frequent bus riders (33% of those retained in summer 2019 did so).

The occasional weekday off-peak cluster saw its retention rate shrink from 22% in 2019 to 8% during the pandemic. The shares staying in-cluster or switching to fairly frequent off-peak rail use, which were the largest "destinations" for this cluster going into summer 2019, did not maintain their clear leads as destinations during the pandemic. Similarly, the collapse of occasional rail commuter retention rates in 2020 also meant the collapse of the share that, in normal years, increased transit usage intensity to become frequent rail commuters. For occasional weekend users, staying in-cluster was still the most common pattern among retained riders; meanwhile there was a decrease in the second most popular destination, switching to fairly frequent weekday off-peak rail.

When we compare each cluster's retained riders from the baseline to the cluster's full pandemic passenger count including new riders, we observe that retained riders are a larger share of the total for the frequent and fairly frequent rider clusters (80% or above). By contrast, in occasional rider groups retained riders are no more than 30% of all cluster members—this may be because occasional riders find it easier to leave the system during COVID, or because they are more likely to use a new smart card when they return to transit (and thus show up to AFC as a new rider). For frequent bus riders, frequent rail commuters, and fairly frequent weekday off-peak rail riders, the share of retained passengers using 1-day passes, 7-day passes, monthly passes, pay-as-you-go, or Perq is similar to that shown in Figure 4-23 for all cluster members. For occasional rider clusters, however, retained riders were nearly all pay-as-you-go, despite the higher presence of various pass types in both the baseline and full pandemic-period rider pool including new entrants. When including these new pandemic passengers, 40% of occasional rail commuters and 20% of occasional weekend riders used passes—higher than the 25% and 6% among retained riders. Notable shares of new riders who entered the T during the pandemic as various sorts of occasional travelers were thus pass users—most commonly, of the 7-day pass.

The user type smart card meta-data indicates that, unsurprisingly, adult full-fare tickets, which contributed the most riders to any one cluster, was the share that shrank the most among retained users. This is especially clear in the frequent bus riders cluster, where 63% of retained riders were adult full-fare, compared to 69-70% in the baseline and in the summer once we included new riders. Separately, students made up 5% of retained frequent bus riders, whereas they were around 1% of all COVID-era riders in this cluster; both figures however were still far below baseline (13%), marking the general disappearance of students from regular transit use. Among fairly frequent weekday off-peak riders, there was a higher share of senior users among the retained (10%) than what was observed once new pandemic riders were added (7%), and more also than the baseline (7.5%). Occasional bus riders saw higher shares of both seniors and TAP riders among the retained (14% and 9%) compared to the figures once new pandemic riders were added (6% and 4%) and the baseline (6%, 2%). In general, then, it appears that reduced fare riders were more likely to be among the retained members of each behavioral cluster, with those who joined during the pandemic tending more towards adult full-fares.

The starkest takeaway from this churn and cluster switching chart is the collapse in frequent rail commuter cluster, despite its historical role as a consistent bulwark source of both ridership and journeys in summer. These riders made limited attempts to keep engaged in transit through switching to a lower frequency usage pattern or time-shifting to off-peak hours to avoid crowded vehicles; this likely reflects the ability of these commuters to switch to remote work. Meanwhile, frequent bus riders again showed they were the most likely to retain old patterns of transit usage; further, fairly frequent off-peak rail riders who stayed in the system tended to continue traveling off-peak (though sometimes less frequently). This again suggests that bus ridership, especially among the most intensive riders, do not pose as much uncertainty to the COVID-19 transit recovery; to a lesser degree fairly frequent weekday off-peak rail riders are similar. Meanwhile, managing the recovery of rail commuters requires facing greater uncertainty, but also provides more room to shape the way which which demand returns to the system.

5.5 Caveats

This chapter assessed COVID-19' impact on the behavioral composition of those riding the T. It also traced baseline period riders through to the subsequent summer to examine how their behavior evolved due to the pandemic. This allowed us to draw conclusions regarding the diverging ways in which users of various MBTA modes, and those who rode during distinct times of the day, were differentially affected. However, there are still major groups of riders who may have relied on transit during COVID-19 missing from this analysis due to our data limitations. First, the lack of cash transactions data and the inherent untraceable nature of cash means that we have no visibility into the ridership patterns of some of the most socio-economically vulnerable riders. Second, smart card data only captures the travel patterns of those already riding transit. It does not capture the travel demand of those who lack transit access, or changes existing users hope to see in the network design or system accessibility to aid them during the pandemic. Additionally, smart card data also does not tell us the *reason* behind the dramatic churn rates observed during the pandemic; we cannot separate, for example, transit riders who switched from rail commuting to driving, from those who switched to remote work.

Surveys and interviews will be needed to give greater clarity on these questions, with sampling focused on churned riders, cash-based riders, and riders with both limited transit access and limited access to other modal alternatives. MBTA Fiscal and Management Control Board's ongoing ridership survey effort, for example, has found that in April-May 2020, 40% of low-income riders (those making less than $43,500 a year) interviewed said they had used the MBTA in the past week, while only 10% did among individuals making over $76,000 a year [40]. By January 2021 with reopening underway, the gap had narrowed with 70% of low-income riders and 35% of those making over $76,000 a year reporting using MBTA services in the past week. This is still far from the pre-pandemic baseline, however, when the shares were roughly equal.

Finally, for comparability across the baseline and COVID-19 time periods, we applied the k-means model trained on baseline 2020 data to tease out behavioral clusters in the pandemic-era ridership ODX data. Future analysis can re-train a new clustering model on ODX data across various stages of the pandemic, and assess how the cluster characteristics, as represented by centroids, may have evolve over time.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 6

MBTA's COVID-19 response and policy recommendations for recovery

The passenger clustering and churn analysis described the evolution of transit ridership under the pandemic. In this chapter, we summarize the lessons from those quantitative chapters and contextualize those results within the COVID-19 situation in Metro Boston and the MBTA response. Then, this chapter compares Metro Boston's situation against that faced by 12 other major transit agencies across the U.S., before briefly looking to international transit agencies to broaden the horizon of the types of innovative responses that agencies brought to bear. From this combination of quantitative analysis and policy comparison, it then draws some overarching lessons for the MBTA's recovery regarding its immediate, medium-term (next two to five years), and long-term (five years or more) actions.

We frame recovery not only as an opportunity to return to pre-pandemic service levels, but to leverage the crisis as a policy window to reconsider how transit delivers on its goals of providing access and sustainable mobility. Substantial uncertainty remains with regards to the trajectory of COVID-19 and the impact this will continue to exert on transit and its user groups, so our recommendations can only be guidelines. However, we provide these in order to contribute to a framework for helping to structure the MBTA's thinking as the agency plans for recovery.

In the short term, the need to bring choice riders especially commuters back onto the system can be an opportunity for the MBTA to more closely engage its employer Perq pass partners for service delivery planning. It also opens a window to discussing both fare product restructuring and service delivery (e.g., changes to frequencies), to improve transit's attractiveness to riders with flexible work schedules post-pandemic. In the medium-term, the MBTA will need to reconsider network design and upgrades to meet the mobility needs of a post-pandemic population. The greater ridership stability of bus riders and mid-day rail riders during COVID-19 means that capital projects targeting these groups can move forward even as uncertainty complicates planning for the return of rail commuters. Further, COVID-19 offers a view into the ridership patterns of transit-dependent passengers with choice riders mostly removed, which can be retroactively dissected to support equity objectives in system planning. Finally longer term, a shift towards a more stable operating revenue mix can enable and complement improved financial management to offer the MBTA greater resilience for the next crisis.

6.1 Key lessons for COVID-19 response and recovery from clustering, churn analysis

There are several key sets of interacting findings gained from the clustering and COVID-19 churn analysis in Chapters 4 and 5. First, there are *clear modal and temporal splits* in rider behavioral clusters, meaning that different rider clusters exert distinctive operational demands on the system in volumes significant enough to make each cluster relevant for operational planning, especially among the multi-day riders. Among these, *frequent and occasional bus riders, alongside fairly frequent weekday off-peak rail riders, are important to track from an equity perspective* because they cover the majority of reduced fare riders and were the most likely to be retained on the system during COVID-19. This could be due to a combination of factors including transit dependence and jobs with specific, fixed shift schedules and no remote work options (e.g. hospital work and home care). Because the two clusters of bus riders are active throughout the day and the fairly frequent weekday off-peak rail riders ride during the midday hours, when the revenue-intensive rush hour commuters are not on the system, these temporally separated behavioral clusters point to the equity implications of retaining or upgrading bus and off-peak weekday rail service during the recovery, especially as ridership for these segments show more continuity than other rail clusters.

On the other hand, frequent rail commuters contributed disproportionately to Perq, which is by itself about 10% of MBTA revenues. These riders also churned the most heavily among multi-day clusters during COVID-19, relative to background levels, bringing a stronger degree of uncertainty about their post-pandemic demand patterns. Heavy employer engagement may be needed to coordinate and incentivize the return of these riders, now potentially with more flexible office work schedules, to return to the system and bring with them the associated fare revenues. Flexible work schedules may also partially disperse demand from peak to off-peak hours, which may synergize and give MBTA further justification to provide more frequent service during off-peak hours both to service the choice riders shifting into this period, and the fairly frequent weekday off-peak rail users who are transit dependent.

Finally, we note that ridership during COVID-19 offers a unique window into the travel patterns of those who are most transit-dependent. Retroactive dissection of ridership from this period can contribute to future planning for equity and resilience.

6.2 MBTA: Pandemic impact and response

6.2.1 The COVID-19 context in Metro Boston

Massachusetts was one of the hardest hit U.S. states in the early months of COVID-19. The state registered its first case on February 1 and by early May had over 86,000 confirmed cases resulting in an estimated 5,700 deaths, giving it the third-highest number of confirmed COVID-19 cases per 100,000 people, behind only New York and New Jersey [71]. Governor Charlie Baker issued a state of emergency by March 10 and established a "command center" to coordinate the public health response four days after [71]. By March 23, non-essential businesses were ordered to close, and businesses that could enable remote work had transitioned its workers out of offices. An executive order requiring face masks in public places, including MBTA vehicles or facilities, was issued on May 6. On May 18, the Governor announced its "safer at home" advisory asking residents to only leave home for essential activities including healthcare, permitted work, worship, shopping, and outdoor activities. At the same time, he issued the state's four-phase reopening plan.

Massachusetts entered Phase I soon after on May 18, 2020, with limited opening of manufacturing, construction, worship, and health centers. On June 22, 2020, Governor Baker announced entry into Phase II, where a broader range of activities and locations like retail, outdoor and indoor dining, and office spaces were phased in at limited capacity. Phase III began on July 7 with the re-introduction of more leisure activity venues like movie theatres and fitness centers. As Figure 6-1 shows, case loads were low in the state during these warmer months, allowing for quick reopening. After a winter roll-back in reopening due to a surge in cases, Governor Baker re-commenced Phase III plans on March 1, with the state entering the first stage of Phase IV (the "New Normal") on March 22, 2021 as vaccination became more widespread. Phase 1 of the state's vaccination rollout also began during December 2020, with priority for healthcare, first response, long-term/assisted living/home-based care workers [72]. Phase 2 began February 1, 2021 and opened vaccinations to seniors, those with multiple medical conditions, those in low-income and affordable senior housing, educators and child care workers, and certain additional worker categories. By April 19, 2021, Phase 3 of the vaccination plan had been implemented, allowing all people aged 16 or older who live, work, or study in Massachusetts to be vaccinated.

Figure 6-1 shows that correspondingly, Google Mobility Reports indicate retail and recreation travel recovered from 54% below the January 3 - February 6, 2020 reference levels at its nadir in April 2021 to only 15% below by March 31, 2021. Travel to workplaces recovered from -59% to -39% over the same time period—still substantially low compared to normal as offices that allow for remote work have not

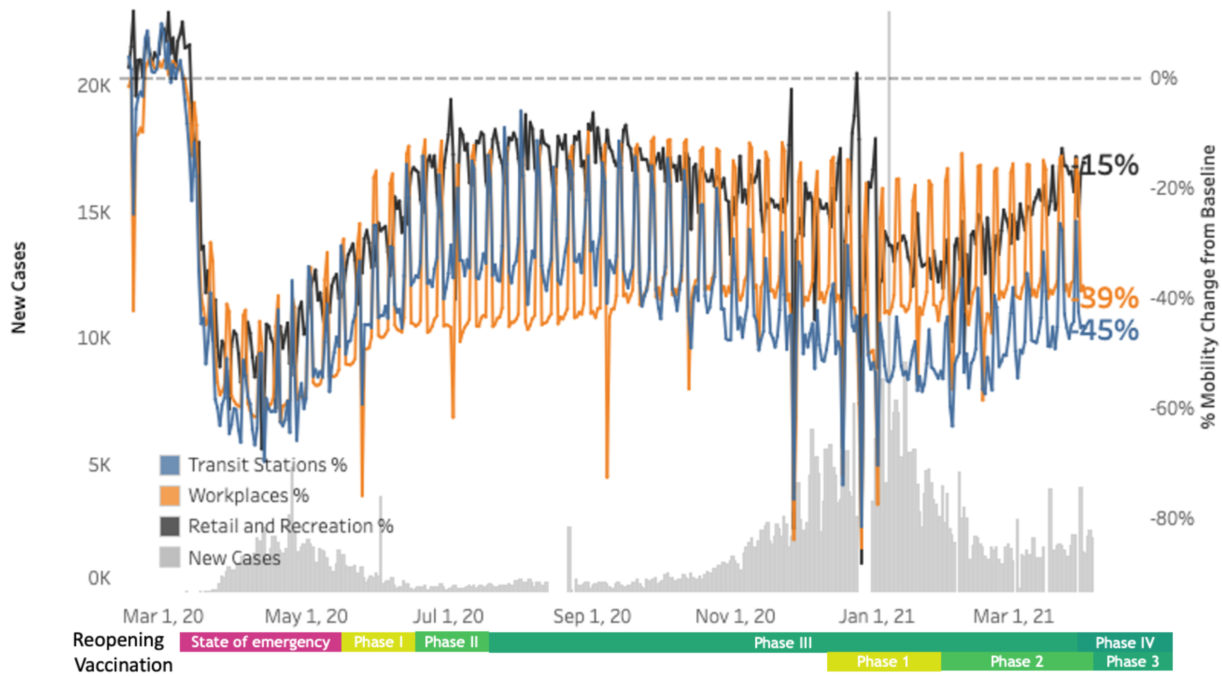


Figure 6-1: Massachusetts daily caseload and mobility patterns

fully transitioned back to in-person. Meanwhile, transit use lagged behind, recovering only from -70% to -45%, potentially linked also to the large role that work commutes contribute to transit ridership (commuting was the trip purpose for 71% of individuals surveyed in the 2015-2017 System-wide Passenger Survey) [4]. Apple's mobility data indicate a similar lag in transit usage recovery for the Boston area in particular—transit was down as much as -81% in April 2020 below a January 13, 2020 reference date and recovered to -27% by April 22, 2021. In contrast, walking and driving have recovered steadily from as low as -66% in April 2020 below the same reference date to 15% *above* for driving and 2% *above* for walking by April 22, 2021 with even higher figures on weekends.[1]

As of April 2021, Suffolk and Middlesex counties, which are the two most encompassing of the Metro Boston area, tallied 208,617 cases among 2.4 million residents. Daily case rates remain relatively high at 18.9 per 100,000 people in Suffolk County and 15.9 per 100,000 in Middlesex County [73]. The share of individuals with at

[1]We use numbers from Google and Apple mobility reports here for transit rather than the MBTA's. This is so that they are comparable to the other data on walking/driving and retail/work travel available from these services, which are not available from the MBTA.

least one vaccine dose reached 47% in Suffolk County and 51% in Middlesex County, facilitating the state's continued Phase 4 reopening [74]. This continued reopening makes more urgent long-term discussions of the recovery trajectory for the MBTA in order to effectively and equitably service the Metro Boston population and continue offering a high-quality sustainable transportation mode.

6.2.2 MBTA response and public reactions

The MBTA's response to the pandemic under its Forge Ahead initiative included enhanced sanitation and public health measures, as well as service cuts for 2021 which were rolled back when new federal relief funding enabled the restoration of pre-pandemic service levels in April 2021. In March 2021, MBTA also released its long-term budget outlook, to address long-term fiscal sustainability after federal funds run dry. These response are summarized in Table 6.1, which the MBTA has proactively communicated to the public via its website, social media, and announcements on MBTA vehicles. Drastic service cuts proposed in November 2020 in particular led to public outcry and independent fiscal evaluations by the Fiscal and Management Control Board, which led to roll-backs to proposed cuts. These in turn were nullified when the MBTA received another $845 million in federal funding on March 30, 2021. These developments highlight not only the continued need to monitor the transit usage patterns of transit-dependent passengers to fulfill access and equity objectives, but the need to continually maintain fiscal health to enhance resilience in the face of unanticipated crises.

Response area	**Measures**
Personal sanitation / health measures	**Masks mandated* by law. Enforced penalties include denial of service, fines up to $300/violation. **Free masks* distributed starting Nov. 2020 in morning, evening peaks at set of key transit stations. **Hand sanitizer* available at stations. **Social distancing* encouraged with decals and messaging.
Facilities, vehicle sanitation / health measures	**Disinfection* of vehicles occurred daily, buses cleaned multiple times/day. High-contact surfaces cleaned every 4 hrs, key facilities every 24 hrs. **Protective barriers* installed to protect vehicle operators; prior to that rear-door boarding was in effect. **Air ventilation* on bus already relatively high with all air on-board refreshed 50-70 sec. Subway air ventilation with HVAC units already recycle and filter all air on-board 14 times per hour, air filter additionally refreshing air every 60 sec. Filters are generally MERV 3-7 changed out once a month, though MBTA is working to upgrade to highest MERV level existing HVAC systems will allow.
Service cuts	**Mar. 2020:* Rail frequency reduced to 9-14 min, buses put on Saturday schedules except for essential routes. **Nov. 2020:* Service reduction proposal released as part of MBTA's Forge Ahead. After public feedback, service reductions scaled back in new Dec. proposal, began rolling out Jan. 2021. **2021 January - March:* Weekend ferry and commuter rail service closed on 7 lines. 20% cuts in frequency on the Green, Orange, Red Lines but only 5% on the Blue which has higher pandemic volumes. 15 bus routes suspended or consolidated. **April 2021:* Service restored to pre-pandemic levels, thanks to more federal rescue funding. **Fares adjusted* to eliminate higher pricing for CharlieTicket v. CharlieCard trips.
Case tracking	None
Recovery planning	**Planning for long-term budget sustainability* after emergency federal funding runs out. **Enhance own-source revenue streams* (real estate, advertising) **Manage debt profile.* **Hedge fuel prices* to minimize unexpected costs. **Manage fiscal controls* to limit expense growth to 2.4%/year for FY2023-2026, below 3.4% annual growth. **Reassess opportunity* for operational efficiencies, reduce cost of professional services.

Table 6.1: Key MBTA COVID-19 responses

Personal sanitation/health measures

As of November 2020 by state and federal law, masks are required for operators and riders on the MBTA network. The MBTA has tied enforcement mechanisms to this, including penalties such as removal of non-compliant passengers from vehicles, denied boarding, and civil fines up to $300 per violation. To facilitate compliance, the MBTA is also distributing free masks during morning and evening rush hours at at Forest Hills, Maverick, Orient Heights, Park Street, Downtown Crossing, Quincy Center, Charles MGH, and Hynes which became a mass vaccination site.

Public transit advocates have voiced concerns about mask mandate enforcement and how the $300 fine would be enforced, noting that this may more directly hit people of color and those unable to afford a $300 fine. Recommendations have revolved around enforcement that emphasizes reminding passengers for compliance, paired with offering free masks [75].

In addition, our analysis in Chapter 5 suggests that distributing masks only during peak times is likely not the most optimal strategy for targeting those who continue to ride during the pandemic. Bus rider clusters and fairly frequent weekday off-peak clusters exhibited more continued ridership than peak-hour commuting clusters, and an aggregated view of pandemic-period ridership also pointed to the rising relative prominence of midday ridership at least as of late summer 2020. Further, the station-oriented nature of these mask giveaways may not be enough for the high share of retained bus riders. The listed stations are already strong in that they cover key intermodal points and points popular among riders of most clusters. However, while Forest Hills is the top popular stop among bus riders (and Maverick is one of the top ones for fairly frequent off-peak rail riders), other major stations used by these high-retention clusters like Nubian or Sullivan Square are missing. To target these continued and transit-dependent riders, who are also more likely to use reduced fare products, MBTA can distribute masks in midday off-peak hours as well as on-board buses along major bus routes. To reduce the labor cost for this expanded range of mask distribution, the MBTA can install small mask dispensers on buses as used by

Atlanta's MARTA and Dallas' DART.

Finally in the personal sanitation category, MBTA offers free hand sanitizer at certain stations and encourages social distancing with messaging and decals. To aid social distancing, the MBTA's app and website provide real-time scheduling and crowding info from automatic passenger counters (APC).

Facilities, vehicle sanitation/health measures

Improved sanitation measures are summarized in Table 6.1. The MBTA notes that air circulation rates in its trains and buses are already better than most office buildings, and that they are attempting to upgrade filters to higher MERV values where possible. The plans have not drawn much critique. In Boston as in other cities, riders consistently cite increased routine cleaning as a factor in coming back onto the transit system [76].

The MBTA's existing air circulation rate is similar to those seen in other city transit systems and on airplanes. However, the MERV level of its filters is low in comparison to the MERV-13 recommended by the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) and EPA for virus protection [77]. It is also low compared to the MERV-8 used by San Francisco's BART, King County Metro, Denver's RTD, and Chicago's CTA. Portland's TriMet uses MERV-10. CTA, Dallas' DART, RTD, TriMet, and D.C.'s WMATA have further adopted electrostatic fogging devices for deep cleaning, with TriMet also adding UV disinfection. Ventilation and sanitation therefore appear to have room for improvement, and given its potential role in bringing back riders may be worth considering as a priority for MBTA's tight recovery budget.

Service cuts & recovery planning

Transit service cuts and its ties to the financial re-planning necessary for recovery have been the most major and most controversial aspect of MBTA's pandemic response. MBTA reduced service frequency in March 2020, with upward adjustments in the Blue Line and essential bus routes to accommodate higher demand [78]. The focus at the time was to discourage mobility to slow the spread of COVID-19 and to protect the health of the MBTA's workforce.

Major service cuts were not proposed until November 2020, when the MBTA cited a projected budget gap of $652 million in FY2022 as justification for scaling back service, alongside data showing the average weekday ridership had fallen from 1.26m trips pre-pandemic to only 330,000 by fall 2020 [79]. The proposal, meant for spring 2021, included cutting subway peak service by 20% on all lines and off-peak a further 20%, as well as early termination for the Green Line E branch. Bus service will no longer run as late into the night, instead stopping at midnight with early service continuing on essential routes. 60 non-essential bus routes will operate 20-30% less frequently; those more heavily used will be down at most 5% and 80 essential routes will not change. 10 more routes were proposed for consolidation and 25 serving under 0.5% of pre-COVID volumes eliminated. These operational changes came on top of a pause in capital projects and reallocation of federal funds to FY2022.

Public response indicated that riders wanted the MBTA to prioritize the time span of its services and access over frequency, aligning schedules with the needs of essential workers and vulnerable communities. Our analysis in Chapter 5 suggests that this would involve re-assessing the scale of the cut to off-peak rail services, which are already less frequent than peak periods, serve a larger share of trips during COVID-19 than during normal years, and are disproportionately used by reduced-fare riders. Off-peak services on bus and train are essential to three of the behavioral clusters that saw the largest continued usage during COVID-19, especially given that bus riders are more likely to use both modes.

The timing of the service changes could also have potentially slowed efforts to prepare for recovery, though admittedly this only became clearer two weeks after the proposal when vaccinations began and reopening re-commenced. Making cuts during a period expected to be the beginning of recovery raised public concerns that services cut will never come back, and questions about why cuts had not awaited further information about the pace of vaccination and federal stimulus funding.

In response to public comments, the MBTA Advisory Board ran an independent analysis finding that the agency's proposal posed long-term risks, and supplying an alternative estimated FY2022 deficit of $528 million which is 20% smaller than the MBTA budget forecasters' worst-case scenario gap [80]. The MBTA then presented an adjusted December 2020 service cut proposal, focusing on transit-critical populations (low-income households, communities of color, disabled, households with no or few cars, seniors). The cuts will eliminate weekend commuter rail services on seven lines, suspend 20 bus routes, and reduce ferry services and bus frequency. They will also reduce subway service by 20% on the Green, Red, and Orange lines and by up to 5% on the Blue Line, which has shown higher ridership levels during the pandemic than other lines. The plan began execution in January 2021 with ferry and commuter rail cuts; in March the subway and bus cuts began with the aim of saving $21m for rest of FY2021.

Immediately after at the end of March 2021, the arrival of $845 million from the federal government enabled the MBTA to restore service to pre-pandemic levels. The MBTA's focus now is instead on planning for longer-term budget sustainability after the expiration of federal dollars through managing its debt, professional service contracts, and fiscal controls. It is also enhancing revenue streams based on its own sources separate from government transfers, including revenues from its real estate assets, parking, and advertising. Such own-source operating revenues had been $101.6 million of its $773 million FY2019 budget. This aspect of revenue mix is a central one towards which transit agencies across the U.S. and internationally have taken varied approaches, which we will briefly review in the two subsequent sections.

6.3 Comparison to other U.S. cities

How did MBTA's initial response to COVID-19 and its subsequent evolution align or diverge from responses taken by other major U.S. transit agencies in its peer group? This section gives an overview of 12 additional major transit agencies in the U.S., excluding New York's MTA which given its scale is arguably too extensive and large to compare with the MBTA. These agencies include 1) Atlanta's MARTA, 2) Chicago's

CTA, 3) Dallas's DART, 4) Denver's RTD, 5) Harris County METRO in the Houston area, 6) Los Angeles' LAMTA, 7) Portland's TriMet, 8) Philadelphia's SEPTA, 9) King County Metro in the Seattle area, 10) Washington D.C.'s WMATA, and the 11) BART and 12) MUNI, both located in the San Francisco Bay Area.

Figure 6-2 presents a heatmap comparing these agencies in terms of 1) system characteristics such as typical ridership volumes, mode split, and revenue sources; 2) COVID-19's impact on the counties in its catchment area, ridership, agency's budgetary situation, and availability of federal emergency funding; 3) agency response to promote personal sanitation, 4) agency responses to improve facilities/vehicle sanitation, and 5) service cuts. We note also that agencies are planning for recovery with some service cuts reversed thanks to federal stimulus money under CARES in 2020 and CRSSA in 2021; however, none have fully restored service to pre-pandemic levels. The heatmap colors are set by row, to indicate how values in one city compares to those observed among its peers. Details about the response for each city are listed in Appendix A.

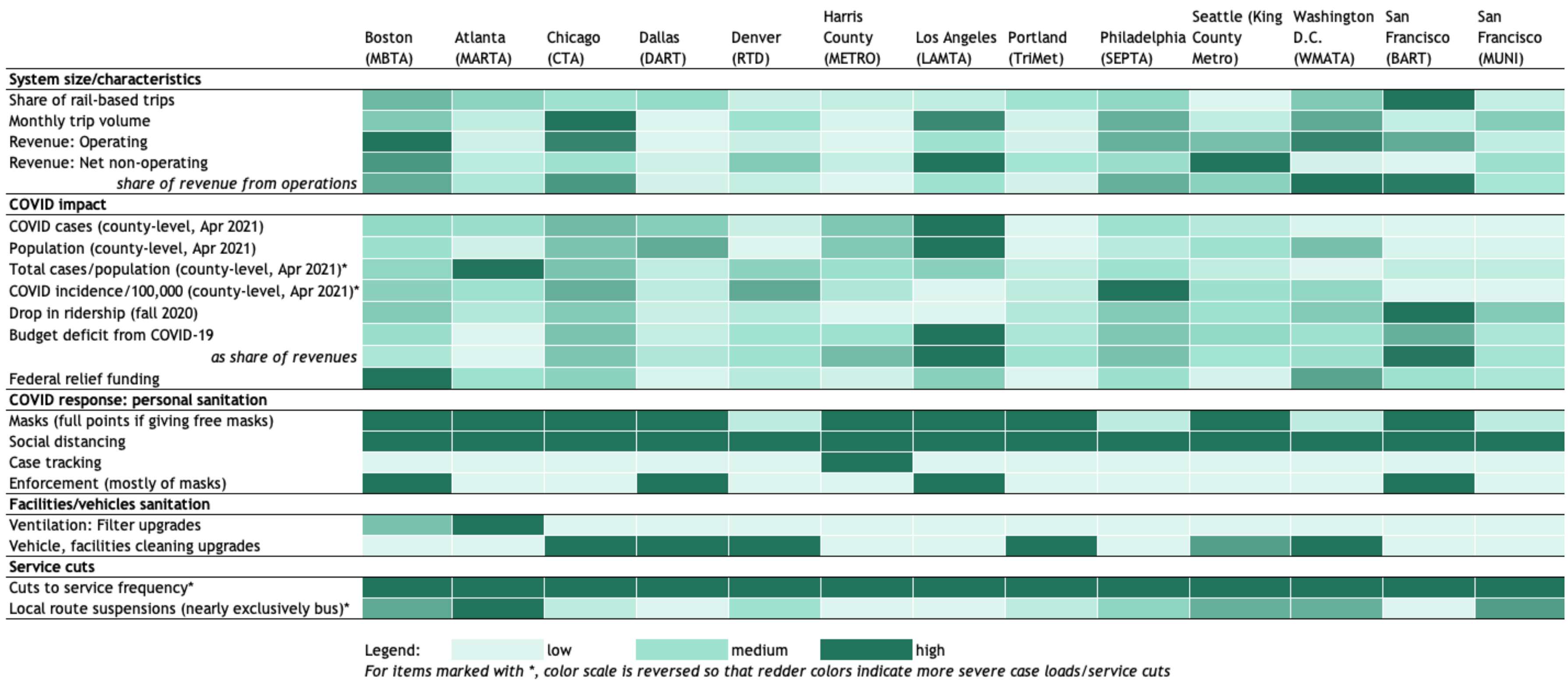


Figure 6-2: Heatmap: COVID-19 impact and response in MBTA's "peer" U.S. transit agencies

The heatmap shows that agencies of a range of modal splits, ridership volumes, and budget sizes were hard-hit by COVID-19 in terms of ridership and finances. LAMTA and Harris County Metro retained more ridership than most, with trip volumes halving. Meanwhile the most-impacted, San Francisco's BART, saw volumes still down nearly 90% even in fall 2020. Because rail services tended to decline in ridership more than bus across the board, BART's poor ridership levels as a purely rail service is not surprising. The MUNI, which has overlapping geographic territory but a combination of trolley and bus lines, experienced closer to a 70% ridership decline, more similar to that experienced by MBTA and CTA.

The sampled agencies also had a diverse array of revenue streams to support day-to-day operations, with younger systems (MARTA, DART, RTD, Harris County METRO, LAMTA, TriMet, King County Metro) generally less dependent on operating revenues. DART, for instance, met 88% of operating expenses from revenue sources not directly affiliated with its operations. Sales taxes were the single largest source, which collapsed by half once the pandemic hit. Harris County Metro, RTD, LAMTA similarly suffered from a collapse in sales tax revenues during the pandemic economic slowdown. WMATA and SEPTA are subsidies-dependent, with SEPTA's subsidies typically tied to operational metrics such as passenger volume. Advertising and parking fees were also often listed among revenue sources.

Yet despite this variety of funding streams, none of the agencies were able to consider restoring full pre-pandemic service prior to the influx of CARES and CRSSA federal relief funding, which totalled between $185 million for the smaller TriMet system to $1.97 billion for the MBTA. The economy-wide reverberations of COVID-19 meant that agencies with more diversified revenue sources outside of fares and other operating revenue sources did not necessarily survive notably better through the crisis. As the MBTA Fiscal and Management Control Board pointed out in its Five-Year Operating Budget Pro Forma (FY22–FY26), careful fiscal management is essential in parallel to having strong sources of revenue.

MARTA was perhaps in the best fiscal shape, with its board putting the credit on careful fiscal management that had led to nine consecutive years of balanced budgets

[81]. This allowed it to set aside \$150 million of its rescue funding for "COVID-related losses" in FY2021 and \$65.6 million in FY2022. Surpluses will be placed in a sales tax reserve, which MARTA expects to reach \$272.5 million by end FY2021; this reserve will be used to cover any potential deficits through 2025 as the economy and ridership recovers. It gave its union employees the previously negotiated 3% raise in FY2021, avoided layoffs, set aside \$20 million in contingency funds for COVID-related personnel expenses, and pledged that it will not increase fares in FY2021. It also continued with its MARTA 2040 expansion program, including bus vehicle upgrades, BRT service, and rail renovations. MARTA did, however, undertake one of the most drastic bus route reductions, although all 110 routes were again active on a reduced schedule as of April 24, 2021 [82]. A year previously on April 20, 2021, MARTA had cut 70 of its 110 routes to divert its full fleet of 540 buses onto 34 essential routes (while adding a circulator and maintaining 6 more routes) to increase social distancing on vehicles along the most-trafficked corridors, and give time for the buses to be outfitted with ionizing air filtration equipment.

MARTA's was one of the largest set of route suspensions of any agency reviewed, and occurred very early on in spring 2020. It drew concern from members of the public regarding the impact on transit-dependent and vulnerable communities. The MBTA, King County Metro, WMATA, MUNI were the only others that had begun to make more than a handful of bus route suspensions; among these, King County Metro announced suspensions in June and MUNI announced in April 2020 as well. MUNI's cuts were the steepest—70 lines were cut and only the 17 most-used remained. This occurred around the time when 40% of operators were out in one week on quarantine [83]. By contrast, all agencies made reductions to frequency and hours of operation. This general approach also aligns with feedback the MBTA received in fall 2020 after its service suspension proposals—riders would rather have less frequent service and longer operating hours, than no service and therefore no transit access at all [80].

Mask-wearing and encouraging social distancing are the most commonly promoted COVID-19 responses, though few places enforced either the mask mandate or its associated penalties. The MBTA, DART, LAMTA, and DART are among the few that do, and all these complemented enforcement with distribution of free masks to lessen the burden on those with more limited mask access. MARTA, CTA, DART, and King County Metro installed mask dispensers on buses, so that riders can obtain masks directly while keeping one's distance from physical contact with a human transit agent. Social distancing measures ranged from cost-effective signage and cordoning off every other vehicle seat, to more expensive measures such as running more frequent buses when necessary to reduce crowding (MBTA, MARTA, RTD, Harris County Metro, BART, switching to longer buses to allow for more social distancing (MUNI, LAMTA), and using automated passenger count data to offer "crowdedness" trackers to riders to encourage to ride during less-busy times. With regards to case tracking, Harris County Metro was the only agency to conduct any sort, in this case temperature checks of those entering METRO facilities or buildings. Thus there seems to be fairly broad agreement on masks and various physical and informational service changes to promote social distancing—and also broad agreement not to heavily enforce or case track.

Sanitation measures that required upgrades to vehicle ventilation and cleaning were less popular, though all agencies increased the frequency of deep cleans using existing equipment for vehicles, facilities, and high-touch surfaces. Nearly all agencies decided not to take major upgrades to its ventilation and air filtration systems, though most have MERV-8 which according to the American Society of Heating, Refrigeration, Air-Conditioning Engineers and the EPA may not be strong enough to block viruses, which require MERV 13. MBTA, having filters of the lower MERV 4-7 types, did invest in upgrading to MERV-8 on vehicles where this is possible, while MARTA installed needlepoint bi-polar ionization capabilities to the air filtration system of all buses in September 2020, to provide fresh air every 75-seconds. Only a handful of agencies used electrostatic fogging to clean surfaces and the air (CTA, DART, RTD, WMATA) or adopted hospital-grade UV disinfection systems (TriMet).

In summary, whether transit agencies were fare-dependent or not, their financial situations were imperiled because of the wide-ranging economic consequences of the pandemic. Sound long-term financial management and a focus on balancing the budget appeared to have allowed MARTA to weather the crisis and stretch out its federal emergency funds more successfully than other agencies. To the extent possible, agencies tried to save funds through reducing service frequency and hours of operation, rather than cutting routes and therefore potentially depriving transit-dependent riders of access. Finally, while masks and social distancing were broadly encouraged, investing the funding for ventilation upgrades and new hardware for deep cleans were less frequently adopted, perhaps due to constraints in operations and financing. The conclusions most applicable for the MBTA are likely those of long-term financial management and greater wariness in cutting bus routes, both of which the agency has come into focus for the agency over the past year. For the purposes of planning more resilient financing, it would be valuable to track the financial recovery of these different agencies to help assess which revenue mixes were most resilient.

6.4 Best practices from international case studies

COVID-19 is a global pandemic, which opens the gates for those researching U.S. transit agencies to look across the globe for best practices in pandemic response and recovery. However, differences in governance, urban mobility culture, transit system structure and age, funding structure, and other factors limit our ability to directly generalize to the MBTA. Therefore, this section simply presents a few high-level takeaways with the goal of broadening the horizon of possible emergency responses considered in this thesis. We draw from a review of COVID-19 responses by transit agencies in Bogota, Hong Kong (MTR and KMB), London, Madrid, Montreal, Paris, Seoul, Sydney, Taipei, and Tokyo.

Sanitation and service frequency cuts implemented across the board

In all international cities reviewed, masking was mandated while the frequency of vehicle and facility deep-cleans were increased. Like in the U.S., most cities abroad focused on intensivity and frequency of cleaning; fewer upgraded sanitation supplies or brought in new technology to enable more thorough cleans. The exceptions were

Transport for London which installed UV disinfection lights, Bogota which upgraded to hospital-standard disinfectants, and Hong Kong which deployed robotic hydrogen peroxide misters. Unlike in the U.S., the addition of free hand sanitizer dispensers in stations was a much more common response overseas.

Social distancing was implemented through the range of tools also seen in the U.S., including capacity limits on vehicles, signs and decals indicating safe distances, apps for tracking vehicle crowdedness, encouraging off-peak travel, switching to bus rear-door boarding, and installing transparent barriers to protect drivers. Service was sometimes added (or added back after initial reductions) to reduce crowding. Sydney added service around the Christmas shopping season, Paris introduced shuttles to connect hospitals to more transportation hubs, and London boosted service along school routes.

Besides these limited cases of service boosts, all agencies cut service frequency and operating hours. However, again like in the U.S., there were very few suspensions of entire routes. Hong Kong's KMB suspended a handful of routes starting March 2020; in some cases stations along a train line were closed to concentrate transit personnel and resources in more essential locations (e.g. sMadrid closed 183 station entrances with under 500 entries a day).

Investing in automation, distanced technologies to protect transit operators, enable service continuation

Several COVID-19 responses observed in other countries reflect attempts to use technology to automate tasks that have higher contagion risk, or to provide greater distance between individuals who have to undertake these types of tasks. Hong Kong's MTR, for example, deployed robots for spraying vaporised hydrogen into transit vehicles and stations, enabling a deeper clean while protecting the health of transit staff and avoiding the need to hire additional workers the way Sydney had done [84].

Mask distribution via vending machines was introduced in Barcelona, Hong Kong, and Taipei, which from a sanitation perspective has a benefit over the mask distribution seen in the U.S. where transit staff are involved, or where a user grabs masks out of a box and therefore could be physically in contact with multiple masks.

Distanced technologies, though not automated, have also been adopted by some cities to run temperature checks and therefore protect both employees and riders. Taipei deployed infrared systems in March 2020 across 16 high-volume stations and people with temperatures over 38 Celsius were refused entry. Bogota used non-contact infrared guns to check temperatures at BRT stations. Hong Kong's KMB similarly used infrared thermometers, but specifically to screen employees; thermal cameras were similar used for employee temperature checks in London and Madrid.

On the more systemic end of the automation spectrum, Paris was able to keep its fully automated Metro lines (Lines 1 and 14) running at 100%, facilitated by the elimination of concerns regarding the safety of in-vehicle operators. In Madrid, 225 transit stations have electronic turnstiles, which were programmed to count smart card taps and lock automatically for 2-3 minute blocks when too many passengers have attempted to enter a station area during a pre-specified time interval, with tighter thresholds set for higher-risk areas.

Capital expenditures like these more systemic improvements, however, require both long-term financial health and political support for transit as an essential service in the urban ecosystem. Hong Kong's MTR is often held up as a model of a financially stable and self-sufficient transit agency; 2020 was the first year it recorded an annual loss since publicly listed as a corporation in 2000. Despite the loss, it has brought service back to nearly 100% pre-pandemic levels and lowered fares by 1.5% to regain demand. Hong Kong is, however, relatively unique in terms of its density and the coordination of land use and transit planning. The MTR's "Rail plus Property" model is also unique, under which the government sells the MTR land rights for the stations and depots along the route and the MTR makes revenue from the development of properties, whose values it has increased through its transit development [85]. However, U.S. cities generally do not have similar governance structure and coordination between land use and transit agencies as Hong Kong.

6.5 Implications for MBTA recovery

The clustering and churn analysis previously discussed in this thesis, combined with this policy review and existing surveys conducted by the MBTA and its non-profit partners, offer a few areas of recommendations for recovery. We divide these recommendations into the immediate, the medium-term (two to five years), and the longer term (five or more years).

The recommendations below view recovery is not only restoring service to pre-pandemic levels, but also taking advantage of a policy window to reconsider how transit delivers on its goals of providing access and sustainable mobility. With the pandemic impact on transit ridership often measured against a pre-pandemic baseline, there is impetus to consider recovery as a return of service to "typical" levels, or of restoring service as demand returns. However, numerous studies already suggest that the post-pandemic transit scene will not be simply to a return to the old normal. Employers that have built up remote work capabilities are, for instance, considering giving employees greater remote work flexibility even long term [11, 10]. The outlook for these affected choice transit riders is therefore uncertain, complicating operational and fiscal planning but also giving the MBTA a window to shape how choice riders engage with transit as a sustainable transportation mode.

Immediate term: Driving ridership and recovering revenue

The most immediate concern for the MBTA is how to plan service provision in the face of four years of projected deficits, alongside uncertain ridership projections and therefore uncertain fare revenues [40]. To do so under their existing mix of revenue sources, they would need to bring back ridership especially among the frequent rail commuters who are revenue-intensive Perq pass holders, whose path back to commuting is still uncertain. The MBTA would need to negotiate with employers on the timing and flexibility of returns to the office and assess how fare product structures may need to evolve to appeal to new flexible work schedules. With all Massachusetts adults now eligible for vaccines, the return to the office is rolling out and the window for the MBTA to shape this transition can close soon if no actions are taken.

A joint survey by A Better City and the City of Boston released November 2020 assessed plans to return to work among office workers, university workers, and hospital staff [76]. It found the majority of respondents plan to return to pre-pandemic commuting choices, although these shares differed by mode and showed the most negative results for transit. While 29% of respondents traveled by subway pre-pandemic, only 16% planned to return afterwards; the figures for bus were 11% and 6%.[2] However, those planning to drive alone soared from 22.9% to 38%, making it the most popular alternative mode considered, although 92% said they had no plans to buy a new vehicle. Biking also rose from 4% to 8%, while 5% planned to continue remote work. Notably though, 45% of those planning to drive alone said they would switch if given free or reduced fare MBTA passes.

Attracting back these riders, who tend to have greater access to alternative modes and remote work, may comprise a key thrust of efforts to restore revenues and prevent mode switch towards cars; however, methods such as lowering Perq pass prices to draw demand would limit the size of the revenue rebound as ridership recovers. Further, from an equity perspective, reductions in fare prices should also be geared towards benefiting lower-income transit dependent riders who do not have corporate subsidies to support them. This suggests that engagement with institutional Perq partners on return-to-work and fare design, and consideration of other funding sources and regulations (e.g. raising parking pricing, promotion of low-emission alternatives like biking etc.) may be necessary for progressing towards both equity and sustainability objectives throughout recovery.

Medium term (2-5 years): Re-design, upgrade network for post-pandemic mobility needs

Consideration of the differing amounts of uncertainty in bus versus rail ridership can help frame discussions regarding the continuation or delay of capital projects, and

[2]The bus usage figures are low for the MBTA as a whole, and could reflect the survey's focus return to work for office, university, and hospital staff, excluding many retail and essential service sectors.

the use of federal funds, for upgrading the network for post-pandemic transit usage patterns. Ridership on buses were more stable during COVID-19 than rail, especially when contrasted against rail use by commuters. The greater certainty about the level and patterns of bus ridership suggest that efforts to prioritize MBTA's limited budget may be enhanced by focusing on bus improvement projects planned using existing bus usage data, for example bus network redesign in the Better Bus Project.

An action that they should *not* take would be to restart bus improvement initiatives from scratch because of anticipated changes to demand due to the pandemic—our COVID period analysis showed substantial consistency in the pattern of bus use among transit-dependent riders. In general, COVID-19 provided information on the "base" levels of ridership on the MBTA system by transit-dependent riders. The patterns revealed during the pandemic can help improve access and equity on the T during and beyond the recovery.

Preparing for upgrades to rail services will require greater monitoring and planning, as rail's demand recovery is more uncertain especially with regards to peak period travel. However, our churn analysis showed greater retention of transit travel patterns among weekday off-peak riders, so this could be a group to prioritize if the MBTA were to target equity as a part of its service redesign. If commuter volumes during peak hours lessen due to more flexible work schedules, this can support shifting trains into the off-peak periods for higher frequency service, enhancing quality for the cluster of rail riders who frequently ride off-peak.

The medium-term is also when the fare product restructuring conversations with employers begun in the immediate term, as well as the contemporary debate on free fares and other fare equity proposals, should be feeding into broader strategies for fare structure. Our analysis is limited in what it can contribute here—whereas a business market segmentation would apply a particular pricing or product differentiation strategy to each segment in this situation, as described in Chapter 2, this is not always possible with passenger segmentation done for transit analysis. Equity is a key consideration for transit and is defined by user type, for example those who qualify for reduced fares as a person who is blind, with a disability, seniors, or students. A transit cluster is relatively homogeneous in terms of its members' behavior, but as we showed in Figure 5-12 each cluster, though potentially more representative of some reduced fare categories than others, never 100% overlaps with a user type. Reduced or free fares are given according to these manually defined categories, which do not completely match our behavioral clusters. Thus our behavioral clusters, though useful for operational planning and for understanding travel patterns on the network, cannot be directly translated into the basis for targeting pricing (e.g. automatically offering a price discount to all people whose smart cards indicate they frequently use the bus would, for instance, benefit mostly people able to pay full adult fares, even though the frequent bus cluster is 31% reduced or free fare users).

Long term (5+ years): Sustainable, resilient financial management to support operations and service

During a crisis like the COVID-19 pandemic, revenue sources collapsed across the board—fares, tax revenue available for transfers, parking fees, real estate values. However, financial resilience during crisis is not only about the immediate budget impact of the crisis and the response, but the state of the actor as it enters into the crisis as well as how it manages its state over the course of the recovery. From this angle, MBTA's examination of own-source revenues growth as ways to boost operating revenues, even as General Manager Steve Poftak signaled that fares will remain integral to MBTA revenues, may be helpful in terms of preparing fiscally for greater resilience and ways to potentially enable other policy items like free transit fares [86]. Moving this forward would require reconsideration of the MBTA's revenue mix for operations, currently 33-35% from fares in a normal year. This farebox recovery ratio is high compared to agencies like MARTA or DART, which have greater infusions of sales or property tax income. MARTA, for instance, relies on fare revenues for only around one-quarter of its operating expenses. As Figure 2-2 shows, other modes like walking and driving are rebounding to higher levels than transit as the economy reopens and the U.S. economic growth rate picks up, bringing with it a restoration of sales, income, and property tax even as transit fare revenues remain highly uncertain on the back of an uncertain rebound in ridership demand. To inform a more stable operational revenue mix, the MBTA would benefit from monitoring how transit agencies with differing levels of reliance on fare revenue recover financially from the pandemic.

Further, continued efforts to improve the structural health of MBTA's financial situation will prepare it for the next crisis as well as enable opportunities to improve service in normal times, including for the most transit-dependent. More stable revenue streams can aid the MBTA's financial planning. Other considerations for enabling greater financial health long-term include reconsiderations of service contracts, which the MBTA is already undertaking with regards to its hired professional services. Labor union contracts can also be another point of potential renegotiation as the MBTA seeks to save on costs. MARTA, which has been able to balance its budget for nine years leading up to the pandemic, is also taking lessons from the pandemic to reassess how it delivers service on essential bus routes versus more peripheral ones. Learning from this example, the MBTA in the long term may have room to consider innovative solutions such as partnerships with ridehailing to reduce the agency's costs for providing service on routes that COVID-19 has revealed to be less essential for the transit-dependent. Such measures for financial health require longer-term planning, but can pay off in terms of leaving MBTA more resilient for the next crisis.

Chapter 7

Conclusion

7.1 Summary of findings and contributions

This thesis leveraged smart card trip data to partition MBTA's riders into clusters based on the temporal and modal patterns of their engagement with the transit network. Further, by using the cluster labels generated by k-means to turn the unsupervised dataset into a supervised one, we were able to apply an optimal decision tree algorithm to interpret the logic behind how the data was partitioned. We found that for the baseline period of January - February 2020, MBTA's entire passenger population can be segmented into seven multi-day rider clusters and three single-day ones. Among multi-day riders, the first major division was along travel mode. The largest cluster contributor to both passengers (11%) and journeys (35%) was frequent bus riders, pointing to the operational significance of this group. Occasional bus riders were second in size by passenger share (10%). These bus-oriented users tend to ride throughout the day and also engage significantly with rail services, suggesting that this group may benefit from consistent service throughout the day and upgrades to both bus and rail services.

Multi-day rail riders, on the other hand, were more finely split over five clusters depending on the intensity of their transit usage and the time of day at which they tend to ride. The largest among rail clusters by passenger share was the fairly frequent weekday off-peak rider cluster (9%), which was also second among rail users in journey count (17%). The frequent rail commuters cluster was small as a share of total passenger (7%), but because of the intensity of their usage made up 21% of all journeys. Further, because these frequent rail commuters (as well as the occasional rail commuters) are almost exclusively traveling during peak hours, they produce concentrated operational strain on the system. Finally, there is a variegated set of occasional rail riders, including the occasional rail commuters just mentioned, occasional weekday off-peak riders, and occasional weekend riders. None of the rail clusters showed heavy intermodal usage of bus, suggesting a relatively strong segregation of these passengers from the other side of the T's network.

Single-day riders, on the other hand, are clearly partitioned into clusters based on their peak usage and weekend share. This created the single-day peak, off-peak, and weekend clusters. These clusters were among the largest in terms of passenger share (10-18% each) but, because of their rare T usage, contributed only a cumulative 5% of journeys. This suggests they are of little operational importance for MBTA planning.

Joining the smart card trip data to the smart card meta-data and a separate MBTA passenger survey, we found that the frequent rail commuter group was not only operationally significant, but also likely revenue significant. Though they are the second largest contributor to journeys and sixth to passengers, they were the largest share of Perq riders (39%), punching far above their weight. On the other hand, frequent bus riders, occasional bus riders, and fairly frequent weekday off-peak rail riders host most of the reduced fare passengers, suggesting that from an equity perspective, supporting bus and off-peak rail service are essential. This appears to be supported by the observation that bus riders also are more likely to be Black/African American and have lower access to cars or alternative travel modes, according to extrapolations from survey data.

With a baseline model in hand, we then turn to tracking the evolution of these clusters with the onset of the pandemic. Because the pandemic period we consider is August - September and therefore a different season than the baseline, we begin by verifying that the k-means model developed is robust across seasons and adjacent years. Having confirmed this, we apply the model to the reduced passenger pool riding in the summer, which contained only 611,327 passengers, far below the 1.55 million in the baseline. After accounting for "background" cluster switching we saw from winter to summer in a normal year like 2019, it appears that the pandemic shifted the shrunken ridership pool and journey counts dramatically away from rail commuters of both the frequent and occasional varieties. Meanwhile, the two bus clusters saw clear upticks in their share of riders and journeys, as did weekend users (both single-day and occasional). Fairly frequent weekday off-peak rail riders saw a more moderate increase.

This suggests that bus service and to some extent off-peak rail service are supporting more transit-dependent workers who could not avoid travel as easily (e.g. through remote work) or shift to other modes. It also suggests that the corporate Perq program's revenue contributions have collapsed alongside rail commuters, and that the MBTA would need to engage directly with employers to construct a rebound in this major revenue stream. Similar lessons emerge when we track only riders who were already on the T pre-COVID through the pandemic era—churn rates were highest among frequent rail riders, who in a normal year were the least likely to churn. Frequent bus and off-peak rail users, on the other hand, were the most likely to be retained during the pandemic.

Overall, the temporal and modal clustering performed in this thesis allowed us to understand the T's ridership in terms of the frequency and modes of their service engagement, at a level of granularity that allows us to draw operational conclusions. Further, smart card meta-data allowed us to tie clusters to prominent user groups and recognize the importance of restoring Perq-using rail commuters to help plug the revenue hole the pandemic has exerted upon the MBTA. Drawing from these lessons and comparing against the COVID-19 response of 12 other U.S. agencies, we find that in the immediate term, the MBTA needs to bring back commuters to support ridership and revenue recovery while channeling workers back to downtown areas to support economic activity. Because of the sea-change that remote work has wrought on commutes, MBTA will likely need to keep in close negotiation with major employers to assess not only how to schedule service to match return-to-office plans, but also to restructure Perq and potentially other fare products to meet more flexible commuting needs. Meanwhile, bus service needs to continue (rather than be scaled back) to provide service for those who are the mostly to keep riding, and to give them access to economic centers to support a resilient recovery.

In the medium term, any reduction in peak commuting travel on rail can free up resources for potentially improving off-peak service frequency, which through the pandemic has shown more consistent transit demand. The consistency in bus ridership over the pandemic also suggests that pre-existing plans for bus upgrades can likely continue, even as the upending of rail demand would require further study to understand and plan for.

Finally, to prepare for future crises, the example of other transit agencies like MARTA in the U.S. and the Hong Kong MTR abroad point to the importance of financial resilience–which would require the MBTA to establish a stable revenue mix and careful fiscal management. Financial resilience for transit supports equity and economic vibrancy in normal times, as well as in crises and recovery. Transit feeds the agglomeration effect of cities and is a lifeline for transit dependent riders with limited modal alternatives. A more financially resilient transit agency is able to offer service through crisis to support the transit dependent and subsequently for connecting residents to opportunities to fuel recovery.

7.2 Directions for future research

Future research following on this thesis can fall into five major categories—improvements to the clustering analysis, spatial dimensions of demand, pandemic and recovery behavior tracking, surveys, and policy research.

On the clustering front, the application of other clustering methods accompanied by greater computational resources may provide more robust results. Those in the literature (Chapter 2) include hierarchical clustering, which can offer an intuitive interpretation via the resulting dendrogram. Gaussian mixture models can also give a more nuanced view of passenger behavioral segmentation by giving the probabilities of cluster membership for each passenger.

This thesis only produces a limited spatial analysis. More work can be done on this front using census data to understand the socio-demographics of ridership and compare this to the findings from the smart-card meta-data, to assess the degree of agreement in socio-demographic analysis from these two approaches. Further, a deeper discussion of how routes and stations were individually impacted by the pandemic can be valuable for operational planning during the recovery, and for understanding the geography of transit-dependent riders at a more granular scale. This can be done by, for example, clustering tap data at the station/stop level instead of at the passenger level, following in the steps of El Mahrsi *et al* (2017) [46].

During the COVID-19 period, future work can re-train a clustering model on the data to assess how centroids and therefore the definition (potentially also the number) of clusters changed. This would require quantitative methods for assessing the "closeness" of clusters, in order for us to interpret how this new clustering relates to the pre-pandemic one. Viallard *et al* (2020) explored methodologies for such tracking of centroid shifts [87]. Additionally, churn, cluster-switching, and re-clustering analysis needs to be conducted for more time periods during the late pandemic and subsequent recovery to track how each behavioral cluster evolved as vaccines rolled out and Boston reopened.

The quantitative work can be supplemented by rider and employer surveys to provide insight on *why* the mobility decisions we observed were made. Surveys can also be used to gauge riders' and employers' outlook for medium-term recovery. Further, ODX and AFC data miss two key rider groups that are central to equity discussions—cash-paying riders who are untrackable in the system, and people who do not have access to transit due to limitations in the current network design. Qualitative surveys specifically targeting these populations can help us flesh out the transit needs of these vulnerable or transit-limited populations during crisis times and during recovery.

Future work on the policy research side can pursue two key avenues: further tying this descriptive analysis into planning for an equitable transit recovery, and consolidating lessons learned from other major transit agencies to determine how to enhance resilience. Throughout the quantitative and policy sections of this thesis, we discuss "equity" as a goal for public transit and use the term to refer to several different concepts. First and most prominently, we assess equity in our cluster profiling by focusing on the travel patterns and needs of reduced fare passengers (i.e., those who are blind, with a disability, seniors, or students). Second, when assessing the pandemic's impact, we discuss equity in terms of COVID-19's differential behavioral effects across clusters. Third, within cluster profiling we discuss equity in terms of the diverging spatial distribution of rider clusters and their shifts in the pandemic. Fourth, when assessing our data limitations, we discuss inequities in data coverage in the context of groups not captured by smart card data, including cash-paying riders and those without easy transit access. Finally, in the pandemic recovery discussion, we touch upon equity in levels of service for groups of transit-dependent versus revenue-generating choice riders. However, because this thesis is focused mostly on describing MBTA ridership patterns, its evolution under COVID-19, and comparisons to other transit agencies, there was no *overarching* argument about how these aspects of transit equity—and perhaps others not mentioned—can be considered systematically during recovery planning. Discussion regarding the fourth and fifth angles on equity listed in particular have much room for development. Future research can approach this behavioral analysis with a cleaner lens on the equity impact of its quantitative results, its data sources, and its operational implications.

Finally, the COVID-19 crisis offers an opportunity to study and learn from the successes and failures of various agencies' approaches to recovery. Tracking recovery of the U.S. cities presented as well as international cases can help to build the knowledge base on how diverse combinations of financial resilience and operational decisions affect ridership recovery, agencies' ability to invest in the future of their networks, and their capacity to deliver on access and sustainable mobility.

Appendix A

Policy Matrix inputs

A.1 Atlanta - MARTA

System description

- MARTA is composed of bus and rail, with ridership about evenly split between the two. In March 2019, bus ridership was 4.7 million for bus and 5.3 million for rail; in February 2020 the figures were lower at 3.7 million and 4.8 million respectively.
- MARTA is funded by a 1% sales tax in Clayton, DeKlab, and Fulton counties, and 1.5% sales tax in Atlanta. Its operating revenues in 2019 were $141m, and net non-operating revenues $688m.
- It has had a balanced budget for 9 years before COVID hit in 2020.

COVID impact

- MARTA's ridership was hit more heavily on the rail side. Bus trip volumes fell 47% from the average weekday in February 2020, as of April 6, 2021. This is worse than last summer when ridership was down only 20-30% but similar to early summer 2020 when ridership was recovering from a low of -62%.
- Rail was down 80% on average in April and May 2020 compared to February

2020, with a slight recovery to -70% or so in summer, then up to -64% on April 6, 2021.

- Federal funding of $632m helped the agency to shore up COVID-related losses in 2020 with $83m, set aside $150m for FY2021 expected losses and $66m for this purpose in FY2022, with surpluses to be put into a sales tax reserve.
- MARTA moved forward with a 3% raise agreed to with its labor union in FY2021.
- MARTA is able to deliver on its previous promise to not increase fares in FY2021 during recovery.

Sanitation/hygiene responses

- MARTA required masks for riders. In July 2020, MARTA began giving away disposable masks to customers at transit stations. Starting in then, all customers and riders were required to wear a face mask or cloth covering over their mouth and nose.
- MARTA added "tissue-like" mask dispensers in buses. As of early September, 459,000 disposable masks have been given away, although MARTA prepared a total of 2 million masks in July. As of April 7, 2021, free masks are still available on all buses.
- Starting April 6, 2021, all buses were disinfected every 24 hours with additional cleaning between select trips.
- In early September, MARTA added antimicrobial air filters in buses to improve ventilation. Needlepoint bi-polar ionization was also added to all buses to provide fresh air every 75-seconds.
- Starting in March, MARTA began implementing "rear-door boarding" to separate passengers from drivers. In early September, MARTA added polycarbonate shields around bus operator cabs and resumed front-door boarding.

- Standing in the aisles of buses was banned and "do not sit" signs were placed in every other seat starting in early April 2020. MARTA also implemented the use of "Bus Full" signs and a dedicated customer hotline to report a full bus and request another bus. However, as of April 2021, all bus seats can be used, though no standing is allowed–buses are limited to full seated capacity. Extra buses will be provided as needed when existing buses are marked full.

Service cuts

- Starting on April 20, 2020 MARTA reduced its bus routes from 110 down to 40 "essential routes" and doubled the amount of buses on the 34 busiest routes to meet the need for maintaining social distance. As of April 7, 2021, MARTA was running buses on 53 essential routes. All 540 buses had been concentrated on a few key routes to support social distancing–meaning lesser run routes had to be eliminated because there were no buses available. This also gave them enough time to upgrade air filtration on buses.

- As of April 24, 2021, all buses were back though they were not on their regular schedule yet. Weekday rail operates from 5 AM to 1 AM, and runs every 15 min from 6 AM to 7 PM—20 min at other times. This isn't too different from normal schedules, which also do not run 24/7, though frequency could be as often as every 10 minutes on weekdays.

- Changes were communicated by the MARTA On the Go app, social media, website, facility announcements.

A.2 Chicago - CTA

System description

- CTA provides bus and rail, with greater bus ridership. In Dec 2019, bus ridership was at 18.6m over 127 routes, 10,715 stops. The system hit 161k miles traveled per day, which is actually less than rail's. In Dec 2019, rail ridership was at 15.9m over 8 routes, 145 stations, and 230,563 miles per day.

- In 2019, fare revenue was around 44% of the operating revenue at $708m, public funding 53% at $884, mainly from sales tax, and the remainder from advertising and non-statutory funding.

COVID impact

- Bus down only around 60% as of summer 2020. Monthly ridership was down to 5.85m in Apr 2020, and up a bit to 7.3m in June 2020
- Rail down about 80% as of summer 2020. Monthly ridership fell to 2.27m in Apr 2020, and up to 2.955m in June 2020.
- $425m received in CARES funding in 2020, upped to $817m by spring 2021.

Sanitation/hygiene responses

- Masks mandated by the state. CTA has distributed 14,000+ "Travel Healthy Kits" containing reusable face masks, hand sanitizer, and a note with healthy riding tips.
- Overnight cleaning of airport line. Transit stations disinfected 4 times a day, and vehicles cleaned before and during service. Applied electro-sprayers for deeper cleans.
- June 2020 COVID era survey showed 40-50% of customers are dissatisfied with various aspects of COVID response vehicle cleanliness.
- No ventilation upgrades. On rail cars, air is filtered more than once a minute through MERV 8 filters, while on bus the air is filtered 50-70 times an hour. This falls short of ASHRAE standards for keeping viral particles at bay.

- Visual indicators and signs installed as guidance for 6 ft distancing (70,000+ signs and decals). Maximum of 22 passengers per rail car, and 15 per 40-ft bus, 22 per 60-ft bus.

Service cuts

- None completely cut except the overnight airport line. Maintained close to normal schedule during pandemic to avoid overcrowding.
- CTA president said that this is financially difficult but they needed to prioritize the people they were serving and their drivers. Extra buses added sometimes where needed to cut crowding.
- As of March 28, 2021, they've doubled bus capacity as part of recovery.

A.3 Dallas - DART

System description

- DART offers rail, bus, commuter rail. In FY2019, the system saw 31m bus rides, 29m light rail rides, 2m commuter rail rides. So, usage is relatively even between bus and light rail.
- In 2019, the operating budget was $544m, while the total expected revenue to cover operating and capex budget was $1bn. This was gained through $628m from sales tax, $85m from operating revenue, $17m from interest income, $90m from formula federal funding, $78m from discretionary federal funding, $91m from long-term debt issuance, $27m from commercial paper issuances, $14m from other operating contributions, $23 from other capital contributions.

COVID impact

- Overall, trip volumes down 55% overall since March, which means there's still more usage than in many other metro areas. DART expected 71m riders in

2020, but now at 48m.

- Bus and rail down relatively similar percentages, at -70% and -64% respectively.
- $229m received in CARES funding in 2020. Sales tax down by about half due to economic closing.

Sanitation/hygiene responses

- Masks required by governor. DART offers a mask dispenser on every bus and train, as well as hand sanitizer. Local police and DART police have power to execute penalties including $250 fines for non-compliance.
- Additional cleaning during day and h2o2 fogger at night. High touch surfaces cleaned with strong chemical, and a hotline opened for those who wish to call in and report unclean vehicles.
- No upgrades to ventilation. Buses already have germicidal air purification system using UV light, silver, intense heat.
- Encouraged 6 feet social distance with limit of 47 passengers per train vehicle, 20 for 40ft bus and 17 for 31ft bus. Pre-booking system available for those who want to avoid crowding.
- Rear door boarding until protective shields installed for operators.

Service cuts

- As of 10/19/2020, 65 routes returned to pre-COVID levels and 48 were on modified weekday schedules (ie, weekday with around 90% of pre-COVID frequency).

A.4 Denver - RTD

System description

- Rail and bus system, with some smaller bus services contracted out to companies like Transdev, First Transit, Denver Transit Partners, Via Mobility Services, MV Transportation, and Eastern Seals Colorado. Denver Transit Operators operate the commuter rail.

- Annual boardings of 95,041,289 (105,823,906 boardings including Free MallRide and Free MetroRide) in 2019. This leans towards bus, which had 59,685,633 boardings. Light rail had 24,585,300.

- Operating revenue is only small part of total operating expenses: $164m compared to $739.7m in the original 2020 fiscal plan. Non-operating revenue is huge, with $996m including $569m from sales tax (just over half). Next largest sources are grant revenue operating and capital, then use tax, investment income.

COVID impact

- In March 2020, there were around 139k trips on average per weekday compared to 348k at this time in 2019.

- $232 million in CARES funding received in 2020. Another similar amount promised under CRSSA in 2021.

Sanitation/hygiene responses

- Follows state mask mandate.

- Capacity limits of 15 passengers per bus and 20 on larger buses, 30 per rail car. Operators can skip stops if they think 6-feet distance cannot be maintained if they onboard more passengers.

- Protective shields installed for drivers.

- Daily cleaning and new electrostatic cleaning devices.
- No ventilation upgrades; vehicles were using MERV 7-8 filters with air recirculated in each carriage in under one minute.

Service cuts

- Instituted cuts across local and regional bus lines, less frequency for most rail lines.
- Added buses on popular routes to help with social distancing. Maintaining social distancing means x2-x3 more buses and operators were needed to carry the same number of customers. RTD thus rescinded layoffs as soon as it was promised CRRSAA federal funding assistance in 2021, in order to make this socially distanced ridership possible.

A.5 Harris County - Metro

System description

- The network is dominated by bus ridership. There were 17m METROrail boardings March 2019-Feb 2020 compared to 59.9m bus.
- Funding is driven by sales tax revenue. $582m was received in sales tax in 2019, and $602m in 2020 original budget. Grants are next: $111m in 2019, $162m budgeted for 2020. Next are bonds ($57m in 2019) and HOT lane revenue ($7m). BRT was started with a $3.5bn bond.

COVID impact

- Bus trip volumes fell 47% February to October 2020. Train trip volumes were down 51% over the same time period.
- CARES funding totalled $289m in 2020.

Sanitation/hygiene responses

- Masks required and offered for free.
- Hand sanitizer installed on vehicles and midday disinfection of vehicles added as of May 2020. Shields for driver's seat.
- Recommended 6 feet social distancing, with seats labeled as unavailable to encourage separation. When vehicles reach 50% capacity, digital sign will advise those waiting to board to wait for next bus.
- Riders submit to and pass a temperature and health screening for COVID symptoms before entering METRO facility or building.

Service cuts

- Agency is asking people to use Metro only for essential trips.
- Local bus routes are on modified schedules all days of week.
- There is reduced service on two of their lines, while increasing services on busier lines to encourage social distancing.
- The red line, which is more used, will operate frequently during weekdays and less so on weekends. Less used train lines (purple, green) are put on more infrequent schedule.

A.6 Los Angeles - LAMTA

System description

- LAMTA offers bus, light rail, heavy rail, bus rapid transit. It also funds 27 local transit agencies, paratransit.
- 265m rides on directly operated buses in 2019, and 12m on privately operated. 93m rides in 2019 on rail. The system is thus heavily bus dominated.

- Vast majority of revenue is from sales tax ($873m from 4 propositions/measures in 2019), then $437m from the Transportation Development Act, $285m from passenger fares. * also for capex: 1185m, and bonds 1408m

COVID impact

- Bus trip volumes fell about 50% 9/2019 to 9/2020. Rail down 42%, which seems very small but is also reflective of how few rail lines there are in LA.
- $1bn received in CARES funding in 2020.

Sanitation/hygiene responses

- Masks mandated and agency aids riders in getting masks.
- Strengthened cleaning at major hubs, cleaned buses and trains once a day with EPA approved materials.
- Hand sanitizer dispensers at major transit stops.
- Ridershp is low enough that crowding is not an issue, and there is no requirement on distancing and bus capacity limits.

Service cuts

- Many bus lines suspended or working on reduced schedule (weekend schedule); busier lines have additional weekday service.
- Over summer 2020, Metro added trips to 95 bus lines and increased the frequency of stops on some heavily-used lines. It also adjusted running times based on faster speeds due to reduced traffic.
- Larger, 60-foot articulated buses will continue to be employed on some lines to the extent possible to reduce crowding.
- No change for the limited train service.

A.7 Portland - TriMet

System description

- MAX light rail and bus are operated by TriMet; Portland and Western Railroad operate a separate WES commuter rail. In 2019 there were 57m bus rides, but only 39m light rail and 0.38m WES. This is a bus-heavy system.
- TriMet is also funded mostly from taxes and bonds, with an extremely low farebox recovery ratio. 43% of revenues are from tax revenue in 2019 ($411m), 21% bonds, 11% federal opex grants, 12% fare revenues.

COVID impact

- Ridership down about 60% as of October 2017.
- $185m in CARES funding in 2020.

Sanitation/hygiene responses

- Free masks given away to meet mask requirement.
- Closed some seats to keep people apart for social distancing (3 feet).
- Capacity limits of 19-24 people on buses, 24-26 per MAX car.
- Safety panels for drivers.
- MAX (light rail) uses a MERV 10 filter that traps particles over 3-10 microns in size, and replaces air every 7.5 minutes with 84/16 recycled to fresh air ratio. Buses have MERV 8 filters (pretty much the same statistics as MERV 10) but do not filter in air from the outside. Buses have every window open that is possible and will not close them no matter the temperature or weather, because that is the only source of fresh air. One type of bus didn't have windows that could open, so they took them off of the road.

- Uses UV lights like in hospitals for disinfection. \$4.4m additional spending on material (non-labor) cost of cleaning efforts as of 5/29/2020.

Service cuts

- 20% service reduction to cut costs as of May 28, 2020. As of April 2021, there was 10% service reduction for buses with only 1 line eliminated—the rest are just less frequent. The MAX runs every 15 min for most of the day, Sunday service for the weekend.

A.8 Philadelphia - SEPTA

System description

- SEPTA runs bus, trolley, subway, commuter rail. In 2019, there were 292.9m unlinked trips. Of these, 138m were bus, 4.5m trackless trolley (like buses). This is a similar volume to the rail modes, which consisted of 24.4m light rail, 51.9m on the MFSE line, 35.5m on the rail line BSS, 3m on NHSL
- In FY2019, 89% of operating revenue from fares. State subsidies make up most of the subsidy income and depends on number of passengers, senior passengers, revenue from vehicle hours and miles. 7.3% "Other income" includes the subsidies, 2% shared ride program, 0.5% investment income.

COVID impact

- Ridership drops are around 70% for both subway and surface transit (mostly bus here). This is a surprisingly even drop across the two modes—with bus dropping more than in many other major metros.
- \$643m in CARES funding in 2020.

Sanitation/hygiene responses

- Mask mandate installed June 2020

- Enhanced cleaning to twice a day in every vehicle, evaluating new equipment for effective cleaning. Only ran buses with plastic seats, for easy cleaning.
- No change to ventilation, which circulates air every 2-3 minutes.
- Maximum occupancies for 40 ft bus is 20 customers, for articulated bus is 30 customers, for routes 204, 310, 311, LUCY Gold/LUCY Green 10 customers, for trolley 25 customers.

Service cuts

- Summer schedule run on most routes, school trip routes suspended, several key rail stations closed.

A.9 King County - Metro

System description

- Metro contracted to operate and maintain Sound Transit's Central Link light rail line and eight of the Sound Transit Express bus routes along with Seattle Streetcar lines owned by City of Seattle. In January 2020, ridership was still equal to that of January 2019 at 400,000–then in February 2020, ridership was actually up y-o-y, from about 380k to 410k.
- 2019-2020 opex totalled $1.9bn. Revenues include 52% from sales tax, 15% from fares.

COVID impact

- Average weekday transit boardings fell from about 400k in January/February to 140,000 by end-August–or a drop of about 65%. Boardings down about 76% as of 9/28/2020 according to the Transit app.
- CARES funding totaled $530m in 2020, plus $1.5m from a public health fund and $276 from the King County Council.

Sanitation/hygiene responses

- Masks required and given to staff. Installed mask dispensers on 102 buses (105 masks per).
- Rear door boarding until partitions installed to protect drivers.
- Safety straps for holding on to, plus seat signs, to support 6-ft distancing between passengers on the bus.
- Capacity limits of 12 people on 40-ft bus, 18 on 60-foot bus.
- Pre-booking system available for those who want to beware of crowding
- MERV 8 filters used in HVAC were installed 2 years before COVID-19 began, so no further upgrades were made.
- Metro maintenance staff begins daily use of Virex spray in backpacks to apply stronger, more comprehensive disinfectant on high touch areas on buses.

Service cuts

- On April 18 2020, Metro enacted schedule reductions (27% fewer service trips than typical weekday service, 15% fewer trips on Saturdays and 4% fewer trips on Sundays) as ridership was 70% below normal. On May 4, Metro added trips on high demand routes for essential workers and trips. On June 2020, route cuts announced. By September 19, 2020, some routes remain cancelled. More evening and weekend services are being added as well for shift workers, east-west connections improvement, and integration with Commuter Rail, and giving more direct service to key destinations.
- Metro is planning to turn the new Route 160 servicing several large suburbs into a permanent RapidLink I route after COVID.
- Opened vanpool to groups of just 2 essential workers to promote distancing.

A.10 Washington, D.C. - WMATA

System description

- FY2019 ridership was 301.8m, with monthly trips between 21m and 26m. Rail was 175m trips, bus 124m, access 2.3m. WMATA is thus rail dominated. It operates a subway, several bus systems, and paratransit.

- System is highly subsidy-dependent, with 43% of funding coming from federal and jurisdictional subsidies.

COVID impact

- As of September 2020, weekday rail ridership was down 87% and weekend 77%. Bus is down 60% weekdays, and only 46% on Saturdays. This is less severe than over summer, when weekend was also as hard hit as weekday. Green line had most continued usage, especially parts in suburbs and Anacostia....end of Orange, Silver lines too.

- CARES funding totalled $546m in 2020.

Sanitation/hygiene responses

- Masks required, WMATA stockpiled 2m by August 2020.

- Electrostatic fogging of all rail cars and buses, including cleaning inaccessible areas like air ducts.

- Closed key stations to prevent tourist trips.

Service cuts

- March 18 saw services reduced, with trains running every 15 minutes and services closing at 11pm everyday. Bus services began operating on a Sunday schedule with some supplemental routes to prevent crowding. On weekends, services are further reduced, with trains running every 30 minutes and with only about 20-30 “essential” bus routes available. By April 6, rail services began closing at 9pm and bus services began closing at 11pm, a direct result of stay-at-home executive orders issued by the Mayor of Washington DC, the Governor of Virginia, and the Governor of Maryland.

- Services were restored starting on August 16 202

A.11 San Francisco - BART, MUNI

System description

- BART is heavy rail, 118m rides in FY 2019.
- MUNI is mostly bus, plus light rail, streetcars, cable cars. Average weekday bus boarding was at 410K as of 3/9/2020.
- BART is fare-dependent for opex pre-pandemic (60%). Passenger revenue $480m, sales tax proceeds $277m, property tax proceeds $51m, then state transit assistance, parking revenue.

COVID impact

- BART: Drop in ridership was more extreme than in other places in the spring, down 89% as of end-June 2020. By 11/23 it was still down 88%.... Starting on March 1, 2020 the ridership was only down 5% from expected. That decline increased to over 90% by March 23. Since that day, ridership has on average been about 10% of expected from the budget.
- MUNI was down 70% on bus boardings.
- As of 9/2020, MUNI’s transit revenue down 93% and parking revenue down 53%.
- CARES funding of $252m in 2020 for BART, $197m for MUNI.

Sanitation/hygiene responses

- Masks required, and all stations have extra face masks for dispensing.
- MUNI Will replace some of the 40 ft buses with 60 ft to allow for more distancing.
- BART gave away personal hold straps to help people distance without having to touch bars. The vehicles already had MERV 8 and filtered air every 70 seconds, so no new procedure was implemented. They also ran long trains all day and put out distance markers to help with social distancing.

Service cuts

- BART running trains all day to allow for more distancing, and added trains during commuting hours when data shows that there are over 30 people on board.
- MUNI replaced Market Street subway with buses as of 3/30/2020. Servicing only 17 most used routes as of 4/2020, with 70 lines cut; this occurred around the time when 40% of operators were out in one week on quarantine.

Note: Data sourced from transit agency websites and press releases.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

[1] American Public Transportation Association [APTA], "Third quarter 2020 ridership," *American Public Transportation Association*, 2020.

[2] L. Liu, H. Miller, and J. Scheff, "The impacts of the COVID-19 pandemic on public transit demand in the United States," *PLoS One*, vol. 15, no. 11, p. E0242476, 2020.

[3] T. A. Ghebreyesus, "WHO Director-General's opening remarks at the media briefing on COVID-19 March 2020," *WHO*, March 2020.

[4] U.S. Center for Disease Control, "Explore human mobility and COVID-19 transmission in your local area," *COVID Data Tracker*, 2021.

[5] E. Badger, "Transit has been battered by coronavirus. What's ahead may be worse," *New York Times*, 2020.

[6] A. De La Garza, "COVID-19 has been apocalyptic for public transit. Will Congress offer more help?," *TIME Magazine*, July 2020.

[7] T. Machemer, "Will mass transit recover from the pandemic?," *Smithsonian Magazine*, July 2020.

[8] J. Harris, "The subways seeded the massive coronavirus epidemic in New York City," *National Bureau of Economic Research Working Paper*, April 2020.

[9] J. McLaren, "Racial disparity in COVID-19 deaths. Seeking economic roots with census data," *NBER Working Paper*, vol. 27407, 2020.

[10] McKinsey, "What's next for remote work," tech. rep., McKinsey & Company, 2020.

[11] D. Altig, J. M. Barrero, N. Bloom, S. Davis, B. Meyer, E. Mihaylov, and N. Parker, "Firms expect working from home to triple," tech. rep., Federal Reserve Bank of Atlanta, 2020.

[12] M.-P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: A literature review," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 557–568, 2011.

[13] A. Alsger, B. Assemi, M. Mesbah, and L. Ferreira, "Validating and improving public transport origin–destination estimation algorithm using smart card fare data," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 490 – 506, 2016.

[14] A.-S. Briand, E. Côme, M. Trepanier, and L. Oukhellou, "Analyzing year-to-year changes in public transport passenger behaviour using smart card data," *Transportation Research Part C*, vol. 79, pp. 274–289, 04 2017.

[15] W. R. Smith, "Product differentiation and market segmentation as alternative marketing strategies," *Journal of Marketing*, vol. 21, no. 1, pp. 3–8, 1956.

[16] M. Wedel and W. Kamakura, *Market segmentation: Conceptual and methodological foundations*. International Series in Quantitative Marketing, 2012.

[17] P. Kotler, *Marketing Management, Millenium Edition*. University of Phoenix, 2001.

[18] C.-Y. Chiu, Y.-F. Chen, I.-T. Kuo, and H. C. Ku, "An intelligent market segmentation system using k-means and particle swarm optimization," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4558–4585, 2009.

[19] C. Jacques, K. Manaugh, and A. El-Geneidy, "Rescuing the captive mode user: an alternative approach to transport market segmentation," *Transportation*, vol. 40, pp. 625–645, 2013.

[20] R. Kuo, L. Ho, and C. Hu, "Integration of self-organizing feature map and k-means algorithm for market segmentation," *Computers & Operations Research*, vol. 29, no. 22, pp. 1475–1493, 2002.

[21] Federal Transit Administration, "Transit and sustainability," tech. rep., Federal Transit Administration, 2021.

[22] J. Berechman, *Public Transit Economics and Deregulation Policy*. North-Holland, 1993.

[23] I. Parry and K. A. Small, "Should urban transit subsidies be reduced?," *The American Economic Review*, vol. 99, no. 4, p. 700–724, 2009.

[24] The Associated Press, "Legislation would support fare-free public transit systems," *WBUR*, 2021.

[25] Massachusetts Rehabilitation Commission, "ADA paratransit service," tech. rep., Massachusetts Rehabilitation Commission, 2021.

[26] A. Basu, "Data-driven customer segmentation and personalized information provision in public transit," Master's thesis, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, June 2018.

[27] L. M. Kieu, A. Bhaskar, and E. Chung, "Passenger segmentation using smart card data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1537–1548, 2015.

[28] M. S. Ghaemi, B. Agard, M. Trépanier, and V. P. Nia, "A visual segmentation method for temporal smart card data," *Transportmetrica A: Transport Science*, vol. 13, no. 5, pp. 381–404, 2017.

[29] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 1 – 12, 2013.

[30] C. Morency, M. Trepanier, and B. Agard, "Analysing the variability of transit users behaviour with smart card data," in *2006 IEEE Intelligent Transportation Systems Conference*, pp. 44–49, 2006.

[31] N. Lathia, C. Smith, J. Froehlich, and L. Capra, "Individuals among commuters: Building personalised transport information services from fare collection systems," *Pervasive and Mobile Computing*, vol. 9, no. 5, pp. 643–664, 2013. Special issue on Pervasive Urban Applications.

[32] M. R. Fissinger, "Behavioral dynamics of public transit ridership in Chicago and impacts of COVID-19," Master's thesis, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, June 2020.

[33] J. Gao, J. Wang, Z. Bian, S. D. Bernardes, Y. Chen, A. Bhattacharyya, S. S. M. Thambiran, K. Ozbay, S. Iyer, and X. J. Ban, "The effects of the COVID-19 pandemic on transportation systems in New York City and Seattle, U.S.A.," *ArXiv*, 2020.

[34] M. Wilbur, A. Ayman, A. Ouyang, R. Kabir, A. Vadali, P. Pugliese, D. Freudberg, A. Laszka, and D. Abhishek, "Impact of COVID-19 on public transit accessibility and ridership," *ArXiv*, 2020.

[35] S. Hu and P. Chen, "Who left riding transit? Examining socioeconomic disparities in impact of COVID-19 on ridership," *Transportation Research Part D*, vol. 90, p. 102654, 2021.

[36] R. Brough, M. Freedman, and D. Phillips, "Understanding socioeconomic disparities in travel behavior during the COVID-19 pandemic," *CPIP Working Paper Series*, vol. 20202, 2020.

[37] K. Sy, M. Martinez, B. Rader, and L. White, "Socioeconomic disparities in subway use and COVID-19 outcomes in New York City," *medRxiv*, 2020.

[38] Commonwealth of Massachusetts, "COVID-19 State of Emergency. Massachusetts Coronavirus Updates and Information," 2020.

[39] N. DeCosta-Klipa, "When will MBTA ridership fully rebound? Likely not for several years, officials say," *Boston.com*, 2021.

[40] A. Gartsman and J. E. Prescott, "MBTA ridership trends and projects (February 22, 2021)"," tech. rep., MBTA, 2021.

[41] Central Transportation Planning Staff of the Boston Region Metropolitan Planning Organization, "SFY2021: Fare Equity Analysis Results," tech. rep., Boston Region Metropolitan Planning Organization, 2020.

[42] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," tech. rep., Stanford University, Palo Alto, California, 2006.

[43] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.

[44] D. Müllner, "Modern hierarchical, agglomerative clustering algorithms," *arXiv*, 2011.

[45] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," pp. 226–231, AAAI Press, 1996.

[46] M. K. El Mahrsi, E. Côme, L. Oukhellou, and M. Verleysen, "Clustering smart card data for urban mobility analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 712–728, 2017.

[47] S. Zhong and J. Ghosh, "A unified framework for model-based clustering," *J. Mach. Learn. Res.*, vol. 4, p. 1001–1037, Dec. 2003.

[48] S. Na, G. Yong, and L. Xumin, "Research on k-means clustering algorithm: An improved k-means clustering algorithm," *Proceedings from the Third International Symposium on Intelligent Information Technology and Security Informatics*, 2010.

[49] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions and Information Theory*, vol. 28, pp. 129–137, 1982.

[50] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[51] M. E. Celebi, H. Kingravi, and P. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," *Expert Systems with Applications*, vol. 40, pp. 200–210, 2013.

[52] D. Bertsimas, A. Orfanoudaki, and H. Wiberg, "Interpretable clustering: an optimization approach," *Machine Learning*, vol. 110, no. 7, pp. 1039–1082, 2019.

[53] H. Dong, M. Wu, X. Dig, L. Chu, L. Jia, Y. Qin, and X. Zhou, "Traffic zone division based big data from mobile phone base stations," *Transportation Research Part C: Emerging Technologies*, vol. 58, no. B, pp. 278–291, 2015.

[54] S. Zhang, Y. Yang, F. Zhen, L. Tashi, and Z. Li, "Understanding the travel behaviors and activity patterns of the vulnerable population using smart card data: An activity space-based approach," *Journal of Transport Geography*, vol. 90, 2020.

[55] L. He, B. Agard, and M. Trepanier, "A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method," *Transportmetrica A: Transportation Science*, vol. 16, no. 1, 2018.

[56] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.

[57] N. Ashidi, M. Isa, S. Salamah, and U. K. Ngah, "Adaptive fuzzy moving k-means clustering algorithm for image segmentation," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 2145–2153, 2009.

[58] P. Bradley, U. Fayyad, and C. Reina, "Scaling clustering algorithms to large databases," *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*, pp. 9–15, 1998.

[59] A. Likas, N. Vlassis, and J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, pp. 451–461, 2003.

[60] M. Breaban and H. Luchian, "A unifying criterion for unsupervised clustering and feature selection," *Pattern Recognition*, vol. 44, no. 4, pp. 854–865, 2011.

[61] R. Varshavsky, A. Gottlieb, M. Linial, and D. Horn, "Novel Unsupervised Feature Filtering of Biological Data," *Bioinformatics*, vol. 22, pp. e507–e513, 07 2006.

[62] G. Goulet-Langlois, H. Koutsopoulos, and J. Zhao, "Inferring patterns in the multi-week activity sequences of public transport users," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 1–16, 03 2016.

[63] D. Bertsimas and J. Dunn, "Optimal classification trees," *Machine Learning*, vol. 106, pp. 1039–1082, 2017.

[64] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[65] Massachusetts Bay Transportation Authority, "2015-17 MBTA Systemwide Passenger Survey," tech. rep., Massachusetts Bay Transportation Authority, 2017.

[66] Massachusetts Bay Transportation Authority, "MBTA FY19 Final Itemized Budget," tech. rep., MBTA Budgets and Financials, 2019.

[67] Boston Planning & Development Agency Research Division, "Neighborhood profiles," tech. rep., Boston Planning & Development Agency, 2017.

[68] MASSDOT, "MassDot GIS," tech. rep., MassDot OpenData, 2020.

[69] City of Boston, "First case of 2019 novel coronavirus confirmed in Boston," *Boston.com*, 2020.

[70] Commonwealth of Massachusetts, "Governor Charlie Baker Orders All Non-Essential Businesses To Cease In Person Operation, Directs the Department of Public Health to Issue Stay at Home Advisory For Two Weeks," *Commonwealth of Massachusetts*, 2020.

[71] Commonwealth of Massachusetts, "Reopening Massachusetts," *Commonwealth of Massachusetts*, 2021.

[72] Commonwealth of Massachusetts, "Phased vaccine distribution plan," *Commonwealth of Massachusetts*, 2021.

[73] COVID Act Now, "County profiles," tech. rep., COVID Act Now, 2021.

[74] Massachusetts Department of Public Health, "Weekly COVID-19 Vaccination Report (April 22, 2021)," tech. rep., Massachusetts Department of Public Health, 2021.

[75] C. Gavin, "Transit advocates have concerns about how the new mask mandate could impact MBTA riders – especially those of color," *Boston.com*, 2021.

[76] A Better City, "Anticipating post-pandemic commute trends in Metro-Boston," tech. rep., City of Boston, 2020.

[77] ASHRAE, "Filtration / disinfection," tech. rep., American Society of Heating, Refrigeration, and Air-Conditioning Engineers, 2021.

[78] Massachusetts Bay Transportation Authority, "MBTA News," *MBTA*, 2020.

[79] Massachusetts Bay Transportation Authority, "Forging Ahead," tech. rep., MBTA, 2020.

[80] Massachusetts Bay Transportation Authority, "Forging Ahead Review," tech. rep., MBTA Advisory Board, 2020.

[81] MARTA, "MARTA balances budget for ninth consecutive year," *MARTA news*, 2020.

[82] MARTA, "All Bus Routes Resume Service April 24," *MARTA news*, 2021.

[83] B. Keeling, "Muni eliminates nearly every route in SF," *Curbed San Francisco*, 2020.

[84] McKinsey, "Metro operator deploys cleaning robot to fight the coronavirus," *McKinsey & Company*, 2020.

[85] R. Cervero and J. Murakami, "Rail and property development in Hong Kong: Experiences and extensions," *Urban Studies*, vol. 46, no. 10, pp. 2019–2043, 2009.

[86] Nik DeCosta-Klipa, "Federal funds are helping the MBTA cover losses due to the pandemic in the short term. After that is the problem," *Boston.com*, 2020.

[87] A. Viallard, M. Trépanier, and C. Morency, "Assessing the evolution of transit user behavior from smart card data," *Transportation Research Record*, vol. 2673, no. 4, pp. 184–194, 2019.