INVESTIGATING A NEW APPROACH TO MARKET CHANGE MEASUREMENT IN PUBLIC TRANSPORT

Prudence Nicole Blake

BRTP (Hons I)

A thesis submitted for the degree of *Doctor of Philosophy* at

Monash University in 2020

Sustainable and Effective Public Transport Graduate Research Industry Partnership (SEPT-GRIP)

Public Transport Research Group (PTRG)

Institute of Transport Studies

Department of Civil Engineering

Copyright notice

Abstract

Growing public transport ridership is a universal goal of cities seeking to address the impacts of growing populations, traffic congestion and climate change. Despite this, measurement of public transport ridership change has been found to have its limitations. Global practice focusses on net increases or decreases in market size neglecting daily fluctuations in new, lost and retained ridership. This limits the scope for understanding market retention.

To fill this research gap, this thesis develops and tests a new framework termed ‘customer fluctuation’. This measures the relationship between new, lost, retained and returning public transport users on a periodic basis. This framework uses an a-priori segmentation approach to aid definition of each market change segment based on travel patterns.

The Thesis develops and tests two alternative approaches to measuring market fluctuation segment sizes applied to a case study in Melbourne Australia. Methods include an approach based on smart card data and an alternative approach using primary surveys.

Smart card data was used to test the customer fluctuation concept by measuring customer fluctuation across all modes of public transport within Melbourne. Initial exploratory testing was conducted to identify the appropriate measurement period and temporal unit for measuring fluctuation. These tests found that using months as the temporal unit and two-years as the measurement period allowed for the most simple and robust measurement of customer fluctuation. Results found that all modes had very high market share of lost users and low share of retained users. These results imply that growth of public transport markets is very sensitive to the net impact of lost vs new users.

The primary survey approach also estimated market fluctuation segments sizes but in addition was also able to explore behavioural factors influencing market change. The survey approach found a high market share of retained users and a low market share of lost users across all public transport modes. These results conflict with those found using the smart card data method. Behavioural evidence from the survey established limited variation in customer fluctuation behaviour between demographic variables (such as age, gender or employment). However, factors linked to customer fluctuation behaviour varied between market change segment. New, lost and returning users identified a range of personal factors motivating behaviour. For returning users, features of the trip taken were linked to their behaviour. Service quality factors, a common focus of public transport operators, were the leading factors influencing retained users.

A comparison of results between the two research methods identified significant variation in estimates of customer fluctuation segment size. The research established a range of limitations evident in both methods; each impacted the estimates of segment size in different ways which in aggregate were found to explain the net differences between estimates made by the different approaches. To improve estimation of market fluctuation segment size, a more robust hybrid

approach is proposed to address the limitations of both survey and smart card data methods. With the refinements proposed, customer fluctuation would be a valuable tool for both understanding public transport markets and developing targeted strategies to better retain and increase the frequency of use by existing public transport users.

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature:

Print Name: Prudence Blake

Date: 1 October 2020

Acknowledgements

I would like to thank all those who supported me in the development of this research thesis. I would like to thank my academic supervisors; Professor Graham Currie, Dr Alexa Delbosc and Dr Zhenliang Ma for all the advice and guidance they have provided for this thesis. This research could not have been completed without their wisdom, guidance and critical thinking while we tested a new concept and kindness in times of despair from that testing.

I would also like to extend by thanks to my Industry Supervisor, Dr Chris Lowe and the support from the Bus Association of Victoria in developing this thesis. The enthusiasm for this work and insights provided were invaluable to this work. Further I would like to thank the industry support received from Transport for Victoria, who generously provided the smart card data that made this research possible.

I am also grateful for the broader academic support that was received. This research was also part of the Sustainable and Effective Public Transport Graduate Research Industry Partnership (SEPT-GRIP). Being able to complete this project with the team of outstanding researchers also involved in this program was an incredible delight and I would like to thank them all for the support, conversations and fun that was had along the way. Thanks in particular to Homayoun, who was a patient teacher and provided much needed finesse to the coding required for this analysis. I would also like to thank James and Taru, who were valuable soundboards and provided me with valuable teaching opportunities during my research.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

I would like to thank my family and friends for their ongoing support and understanding during the completion of this thesis. To my family, my partner Will and all those that offered support during this thesis, I am grateful beyond words. Thanks to Lily Cornish, Matilda Knowles, Marli Roberts, Max and Ange Hardy, Josephine Zavaglia, Michael Horn, Matthew McCartney, Darcy Fleming and Leo Roberts – all of you played a significant role at different times during this journey and I am indebted to each one of you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CLV	Customer Lifetime Value
DBSCAN	Density-based spatial clustering of applications with noise
PTV	Public Transport Victoria
TFV	Transport for Victoria
VISTA	Victorian Integrated Survey of Travel and Activity (Household travel survey)

List of Definitions

This thesis uses several terms from the field of marketing as well as purpose built terms that are not commonly used in the transportation and engineering fields, and as such require clear definitions for this work. The following list defines various terms that are often used in this thesis, as they are meant to be understood in this context.

Contractual (market) setting	A market where there is a contractual agreement between the amount of use and retention as well as inbuilt time periods for contract renewal (Ascarza and Hardie, 2013) e.g. telecommunications.
Customer churn	As defined by Kamakura et al. (2005) the 'tendency for customers to defect or cease business with a company'. The relationship between defecting (lost) customers and retained (existing customers) is the primary component measured by a customer churn approach. The term 'churn' is used interchangeably in the literature with 'attrition' (Libai et al., 2009), 'defections' (Reichheld and Sasser, 1990, Neslin et al., 2006), and 'turnover'.
Customer fluctuation	A measurement concept developed in this thesis. The concept seeks to measure the market change in time by segmenting markets into disaggregate groups based on ridership change. This concept measures the interplay between new, lost, retained and returning customers within public transport markets.
Lost user	For measuring customer fluctuation, a 'lost user' is a user that travels within the measurement period (either starting or continuing travel) and then ceases travel for 2 or more consecutive months and the remainder of the measurement period.
Measurement Period	The length of time identified for the measurement of a concept as set by a start date and end date.
New user	For measuring customer fluctuation, a 'new user' is a user that commences using a service after the start of the measurement period and continues using that service regularly (see retained user) for the remainder of the measurement period.

Non-contractual market setting	A market where there is no contractual agreement between the provider and customer and thus no obligation for a customer to repurchase or reuse a service (Reinartz and Kumar, 2003). Non-contractual settings include retail markets or public transport markets
Public transport markets	This term is used to describe all people that have paid to use, might use or continue to use public transport services. In this thesis, it is used to refer to the public transport market available within metro Melbourne.
Retained user	For measuring customer fluctuation, a 'retained user' is a user that consistently travels on a public transport mode for the duration of the measurement period with no more than 25% of the total study period.
Returning user	For measuring customer fluctuation, a 'returning user' is any user that rides sporadically in the period, consistently starting and stopping travel.
Seasonality	seasonality seeks to measure variations in transport ridership as they can be linked to different temporal factors or seasons (Bocker et al., 2013). Studies that measure the changes in travel patterns over the course of the year will by nature be impacted by variations in seasonal use of public transport.
Temporal Unit of Measurement	The unit of time (for example months or weeks) that is used to take a measurement within the measurement period.

Part A: Introduction, background and approach

Chapter 1: Introduction

1.1 Overview

This thesis investigates a potential new approach to market change measurement in public transport. The approach is developed from a critique of conventional approaches to market measurement based on the disciplines of marketing and public transport studies. The critique, presented in the thesis identifies that conventional approaches to measuring market change in public transport overly rely on aggregate analysis and ignore the disaggregate impact of individual market change segments. Conversely, marketing studies of customer churn, focus on the interplay between some but not all disaggregate segments of change (for example, the relationship between lost and retained customers). These observations have informed the development of a new approach, which forms the topic of this thesis. This sought to measure the change in public transport markets using the separation of ridership change into four change segments; new, lost, retained and returning users. This new approach has been termed '**customer fluctuation**' for this thesis.

Figure 3.5 illustrates the way that new, lost, retained and returning users interact within a market to create customer fluctuation.

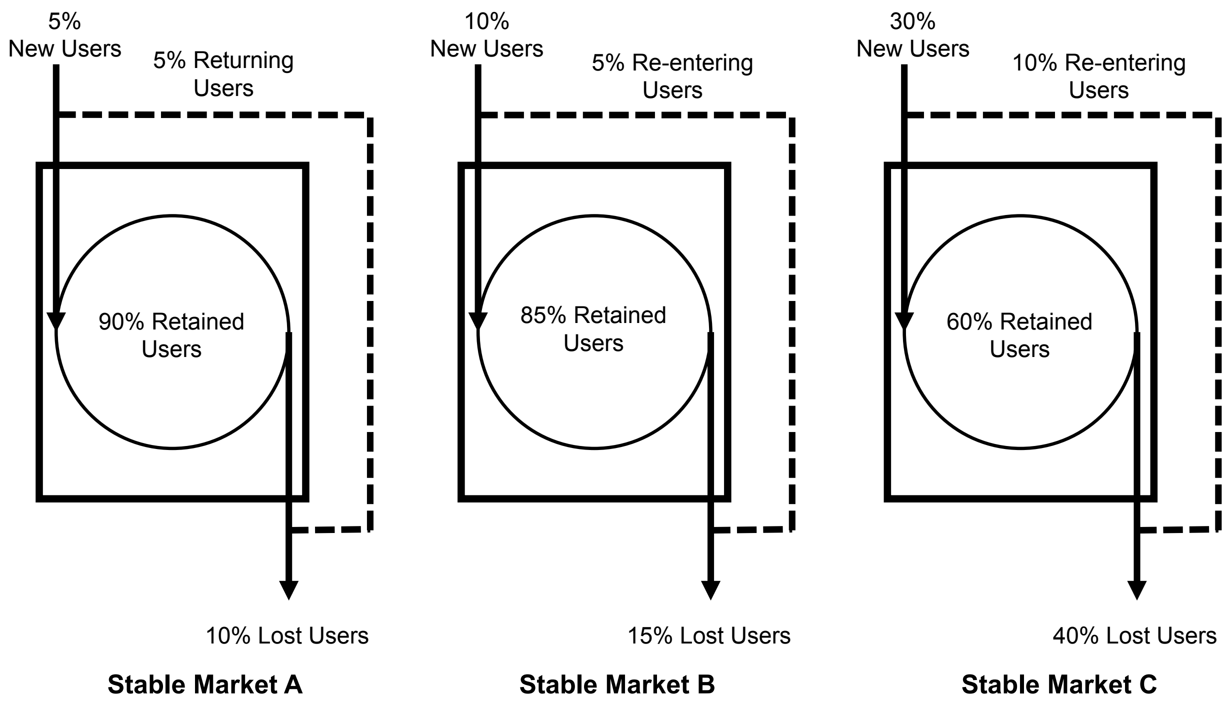

Figure 1.1 - Customer fluctuation within a Stable Market Context adapted by the author from Blythe (2009)

Based on current methods of measurement, all three of these markets pictured in Figure 3.5 would be considered the same i.e. stable markets exhibiting no growth or decline. However, each of these markets are different, becoming increasingly more porous. This thesis seeks to explore different approaches for measurement of ‘customer fluctuation’ with a focus on public transport markets in Melbourne, Australia as a case study. The results of this measurement approach are then used to assess the practicalities and potential of this concept for wider application by researchers and public transport operators.

This chapter is the Introduction to the thesis and outlines the background and motivation for this research, followed by a presentation of the research aims and objectives. The research contributions are then outlined followed by a summary of the thesis structure.

1.2 Background and motivation

Measuring public transport markets is of interest as growing public transport ridership is a fundamental goal of transport organisations worldwide (Currie and Wallis, 2008, Taylor and Fink, 2013, Krizek and El-Geneidy, 2007). The focus for market growth has typically been on attracting new public transport users even though investment in new ridership is costly and takes an extended period to see significant results (Taylor, 2007, Abou-Zeid and Ben-Akiva, 2012, Abou-Zeid et al., 2012, Matthies et al., 2006). Research in marketing suggests that retaining existing customers is a more cost effective and time efficient approach than trying to attract new customers (Reichheld and Sasser, 1990, Reichheld, 1996). In marketing, the ability of a market to retain users is measured using the concept of ‘customer churn'. Customer churn is the tendency for customers to cease business with a company (Kamakura et al., 2005).

There are very few studies that measure the occurrence of customer churn within public transport markets; the only direct application was undertaken on British rail markets by Mason et al (2011). Further, little attention is given to the routine fluctuations of individual patterns of public transport use. The number of transit riders that are being gained (new users) and lost remains unknown in favour of reporting on the aggregate net effects (market growth, stability or decline). This means that although we know the overall changes in a market, we know little about the changing behaviours of the individuals within it. Further the existing measurement approaches over-estimate the stability of users travel patterns, and fail to capture those that might use public transport sporadically (returning users).

It is acknowledged that there is a large volume of alternative tools and approaches currently used to measure and understand public transport markets. Though this does not negate the opportunity to identify new tools, any new approach should offer an advantage over existing methods. In particular, customer fluctuation might be valuable as:

- A more efficient tool for understand public transport markets and internal changes in ridership not currently measured (in the case of Melbourne);

- The identification of new insights regarding ridership and trends in travel behaviour. Though not investigated in this initial exploration, it is considered the impact of this may be assisted if applying customer fluctuation to different public transport contexts and in measuring the impact of the COVID-19 pandemic on ridership.
- A tool to understand markets and any direct linkages between the market fluctuation segments and service modification or marketing/communication strategies that may be used to increase public transport ridership.

It is identified that these factors outline the long-term potential of customer fluctuation as a measurement tool and may not all be achieved through this thesis that focuses on creating and testing the new concept.

Overall in this thesis, a proposed new concept termed 'customer fluctuation' will be used to better understand public transport markets with a focus on aiding public transport operators in growing and understanding market change.

1.3 Research Aims

There are two research aims that guide this project. These aims are as follows:

I. To develop, measure and apply a new concept for market change analysis based on change segments

II. To explore the practicalities and potential of this approach as a means of improving market change analysis

1.4 Research Objectives

To achieve the main research aims, five key research objectives are proposed:

RO1. To understand conventional measures of market change in the fields of marketing and public transport and explore the benefits and drawbacks of these methods

RO2. To develop a new concept for market change analysis based on market change segments

RO3. To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne

RO4. To explore behavioural factors influencing market change segments using survey data for metro Melbourne

RO5. To assess the potential of the new approach for application in the industry

Research objective one aims to provide a deeper understanding of what currently is and isn't being measured within public transport markets to understand market change. This objective also seeks to understand the strengths and weaknesses of relevant approaches to measuring marketing change within the field of marketing.

Research objective two focuses on the development of a new segment based approach for the measurement of market change. The concept uses segments based on changes to individual ridership (for example, starting, stopping or continuing to use public transport) to better understand the disaggregate market change occurring within public transport markets. Concept development includes the refinement of market change segments and setting the initial temporal parameters and approach for method testing.

Research objective three concentrates on the application of customer fluctuation as a measurement tool for public transport markets. This includes an investigation of how the impacts of the temporal unit and measurement period impact the estimated size of market change segments (for example; new, lost or retained users). This also allows for the comparison of different data sources (smart card and survey data) both using a Melbourne case study and their impact on customer fluctuation results.

Research objective four explores the association between customer fluctuation segments and factors influencing travel decisions. These factors are refined from existing studies of factors that influence decisions around public transport behaviour and include factors both endogenous and exogenous to the operation of public transport. This could provide insights into what promotions or targets could be made to promote public transport use.

The final research objective (five), focuses on reviewing the findings of this thesis and assessing the practicalities and potential of this new approach (customer fluctuation) for measuring the market change in public transport markets. This includes a review of the strengths and limitations discovered in this initial exploration of the concept and suggestions for further refinement as relevant.

1.5 Research Contributions

This research provides both theoretical and practical contributions to the existing knowledge of market change in public transport markets. The research contributes to the expanding body of literature regarding the measurement of public transport markets and adapts existing marketing concepts for the creation of a new measurement concept. The development of a new measurement concept includes the identification of parameters that create market change segments based on behaviour change elements and existing research. The research also includes an investigation of the impact of the temporal unit and measurement period on the measurement of market change.

The research also provides an analysis of factors that influence public transport behaviour (both endogenous and exogenous) and how they influence individuals with different travel patterns.

A summary of the key contributions provided by this thesis is as follows:

- The identification and application of the four elements of churn to public transport markets as a new approach for measuring market change through a single measurement approach;
- The development of a mixed method approach and the associated rules and temporal definitions to test the measurement of customer fluctuation – a new method for measuring market change;
- Identification of inherent limitations with smart card data that cannot be addressed and limit smart card data's potential as an independent data source for measuring market change;
- There are significant variations in change segment measurement between measurement approaches using smart card data and primary survey data. The research identified a need to develop a more integrated approach capturing the benefits of both approaches;
- Evidence on the common reasons influencing customer fluctuation behaviour within public transport markets;
- The identification that ongoing fluctuations in market change segments are difficult to measure in practice and may be too complex and costly to implement without further research and testing of new approaches; and
- Recommendations for the development of a new integrated mixed methods approach to reduce the data and technical limitations of both smart card and survey data measurement methods.

1.6 Thesis structure

Figure 1.2 shows the structure of the thesis. The thesis consists of three parts, which are:

- Part A: Introduction, background and approach;
- Part B: Concept testing; and
- Part C: Concept review and conclusions

Part A introduces the thesis, the background and the research approach. It consists of three chapters. Following this chapter (Chapter 1: Introduction) which introduces the topic, Chapter 2: Literature Review outlines the research literature relating to the marketing theory of customer churn and the measurement of public transport markets. The final chapter of this section, Chapter 3:

Framework Development and Research Approach, develops the concept of customer fluctuation and describes the chosen research approach for measuring and exploring customer fluctuation in public transport markets.

Part B 'Concept testing' consists of three chapters, detailing the mixed method approaches used to explore the concept of customer fluctuation as it can be applied to public transport markets. Chapter 4: Measuring Customer Fluctuation using Secondary Data, Part One includes the first test of the customer fluctuation concept using one year of smart card data. This chapter outlines the initial approach, the results found and the conclusions to be used for additional refinement. This approach is then revisited in Chapter 5: Measuring Customer Fluctuation using Secondary Data, Part Two, which applies a refined measurement approach to two years of smart card data exclusively for the bus. Chapter 6: Measuring Customer Fluctuation using a Cross Sectional Survey reviews a cross-sectional survey tool for measuring and identifying the reasons behind customer fluctuation. The appropriateness of these measurement tolls is than discussed in Part C.

The final section, Part C reviews the concepts that have been presented and provides conclusion. This Part consists of two chapters. Chapter 7: Comparison of Measurement Approaches provides a comparison of the approaches used and suggests an improved approach for future research. This is followed by Chapter 8: Conclusions which provides the overall findings and conclusions of this work.

This concludes Chapter 1, which has discussed the background, aims and objectives, contributions to research and structure of this thesis. The next chapter provides the literature review which underpins this work.

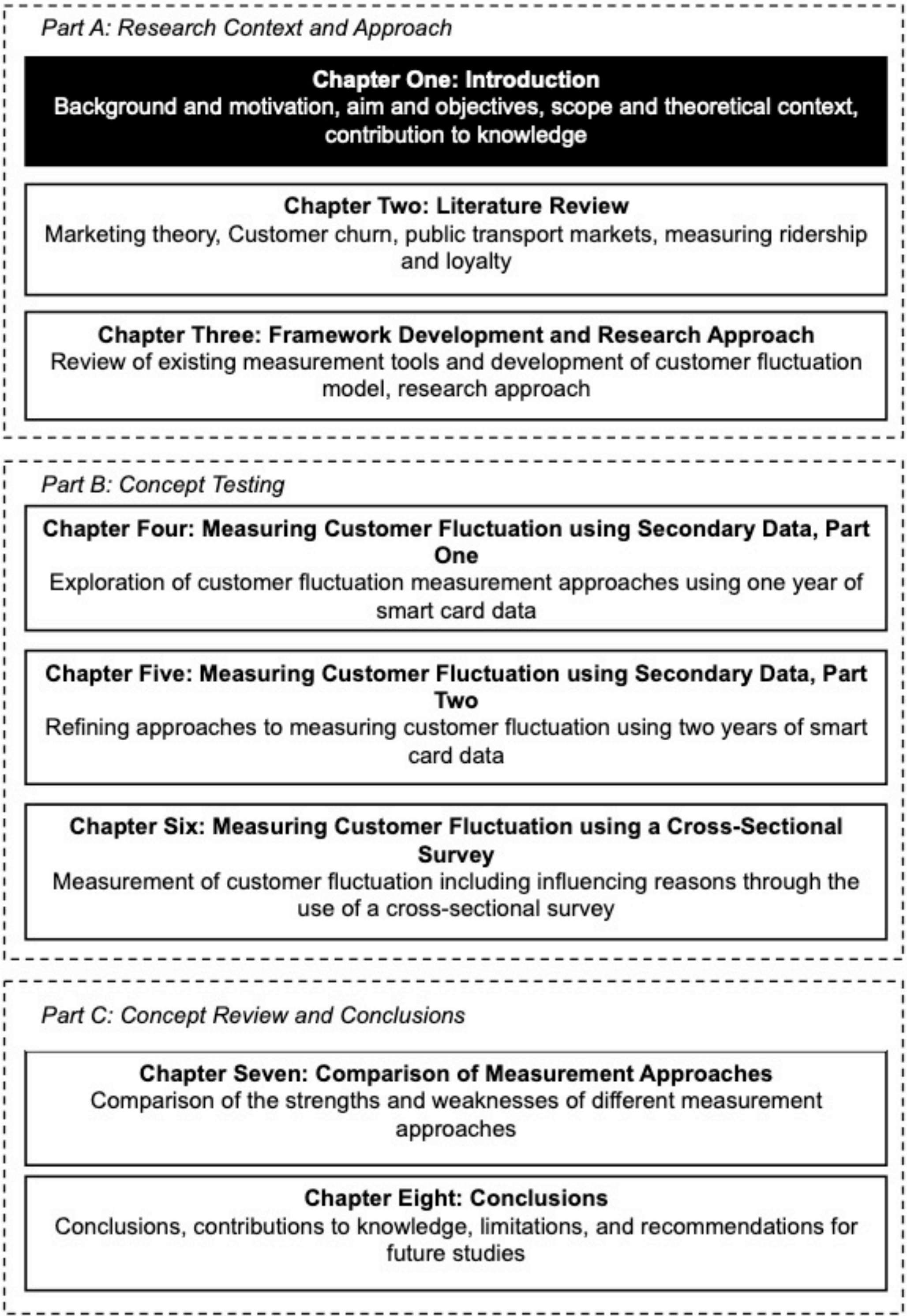

Figure 1.2 Thesis structure

Chapter 2: Literature Review

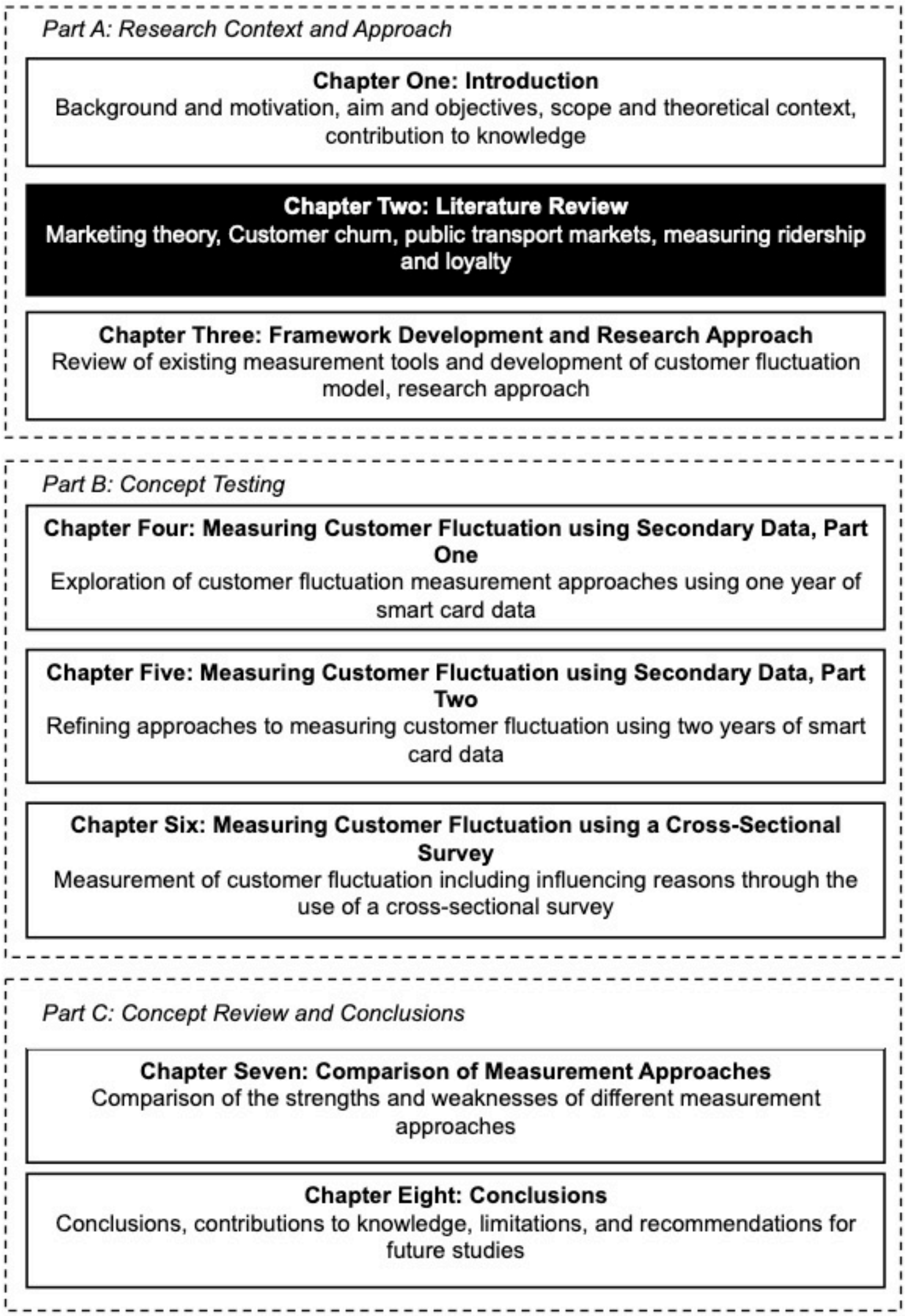

Figure 2.1 - Position of Chapter 2 in thesis structure

2.1 Introduction

Chapter 1 explained the rationale for investigating new approaches to measuring market change within public transport markers. It identified the two-overarching aims of this project, most relevant to this chapter is the first research aim which seeks to develop, measure and apply a new concept for market change analysis based on change segments. This chapter seeks to address research objective 1: to understand conventional measures of market change and explore the benefits and drawbacks of these methods.

This chapter further explores the context of the research topic by reviewing existing research in both the marketing and public transport sectors. As this project seeks to develop a new concept to be applied to public transport markets based on marketing principles the literature review will use an exploratory approach. A focus has been placed on identifying the articles that are most relevant to the development of this concept and the measurement of patterns in transit use behaviour.

A version of this literature review was published as the following conference paper:

> Blake, P., Currie, G., Delbosc, A. & Lowe, C. 2017. Customer Churn: The Missing Link in Public Transport Marketing. Australasian Transport Research Forum 2017 Proceedings. Auckland.

In the next section, the focus is on reviewing the marketing literature on customer churn (2.2). This section will begin with a brief review of the existing understanding of customer churn in marketing. Then reviewing the existing marketing theories and finally identifying the strengths and limitations and the potential for application to public transport markets.

This is followed by a section on understanding public transport markets (2.3). This section reviews existing attempts to measure churn in public transport markets, followed by an outline review of literature relevant to the measurement and understanding of ridership. The section concludes with a review of the strengths and weaknesses of the identified approaches for measuring market change in public transport.

Section 2.4 Conclusions and Implications for Further Studies, brings together the information and provides recommendations for the development of a new measurement approach; ‘customer fluctuation’.

2.2 Customer Churn in Marketing

This section reviews the development and understanding of customer churn as it is currently applied in marketing. It begins by explaining customer churn and the associated concepts including customer lifetime value (CLV) followed by the application in contractual and non-contractual settings, customer switching costs and what is termed the ‘leaky bucket’ theory. This section concludes with a review of the strengths and limitations of customer churn as a measurement approach.

2.2.1 Understanding Customer Churn

Customer churn is defined in marketing literature as the ‘tendency for customers to defect or cease business with a company’ (Kamakura et al., 2005) and remains a prominent research area in the field of marketing. The term ‘churn’ is used interchangeably in the literature with ‘attrition’ (Libai et al., 2009), ‘defections’ (Reichheld and Sasser, 1990, Neslin et al., 2006), and ‘turnover’ (Schneider and Bowen, 1985). The relationship between defecting customers and retained (existing customers) is the primary component measured by customer churn.

Understanding the natural level of customer churn occurring in a market is beneficial as it allows us to set benchmark rates of acquisition and defection within an industry or market. This information can be used to identify what degree of observed customer defections (lost customers) is unusual (Riebe et al., 2014). This is useful information for service providers, as an unusually high rate of defections (above the benchmark where available), may be an indication of customer dissatisfaction with current service provision. It can in some cases also allow for comparisons to be made with competitors to identify competitive advantage or disadvantage.

Customer churn has primarily been applied to businesses within a contractual setting, where there is a connection between the amount of use and retention as well as inbuilt time periods for contract renewal (Ascarza and Hardie, 2013). The measurement period in contractual settings can be easily identified by existing time frames for re-subscription or renewal. However, there have been numerous uses of churn in non-contractual settings such as retail markets (Buckinx and Van den Poel, 2005), pre-paid telecommunications services (Tamaddoni et al., 2010) and in a notable case, public transport (Mason et al., 2011).

The measurement of customer churn in its most simple form is the proportion of lost customers to the total number of customers within an identified time frame. Despite its ubiquitous use within the marketing literature, there are limited academic studies available which provide practical guidance for the measurement of churn. There are however, several industry white papers and blogs that identify the common issues with churn measurement. Brady (2014) identifies one of the key issues with measuring churn which is identifying the relevant variables for use. For example, if you look at customer churn as a measurement of customers from the beginning of the study period or of customers at the end of the study period you will arrive at two different results. Due to the complexities of measuring churn, many studies include methods for the validation of churn prediction. For example, Ballings et al. (2012) split their active customers at the end of the explanatory period with 75% used for estimation and 25% used for validation.

Many industries have moved beyond measuring the simple rate of churn to developing complex processes of churn prediction. The theory is that if you can predict a customer that is likely to churn (or defect) you can better target them for churn management and hopefully, retention. This is otherwise known as identifying the 'paths to death' and creating an understanding of when customers leave (Ascarza and Hardie, 2013). Churn prediction is based on several factors including customer history, lifetime predictions (Ballings and Van den Poel, 2012, Glady et al., 2009) and attitude (Bellman et al., 2010).

An issue with customer churn is that it only provides a limited amount of information in regards to market function by focusing on a single relationship between lost and retained customers (Kamakura et al., 2005, Tamaddoni et al., 2014). Yet, all consumer markets experience natural patterns of customer acquisition (gaining new customers) and customer defections over time (Riebe et al., 2014) as well as having a unique percentage of 'loyal' customers (customer retention). All three elements should be measured to develop an accurate understanding of market change. This leads us to review another marketing theory known as "the leaky bucket theory".

2.2.2 The Leaky Bucket Theory

Based on an analogy, the 'leaky bucket' theory is well known in marketing. The analogy asserts that attempting to grow a market without addressing the 'leaks' or flows out of the market is as futile as trying to fill a leaking bucket with water (Morgan, 1992). In the case of applying this to companies, the "leaks" are identified as the percentage of disloyal customers that need to be replaced (Dowling and Uncles, 1997, Ehrenberg, 1972). In industries where defections are routinely measured it can set the minimum requirements for acquisition. For example advertising agencies must have annual new-business that are equivalent to or more than 10% of total billings to replace the 10% of billings that leak each year (Morgan, 1992).

As an example, Figure 2.2 illustrates the internal changes that might be occurring in three different markets that would all be identified as ‘stable’ (no net change) when relying on aggregate measures.

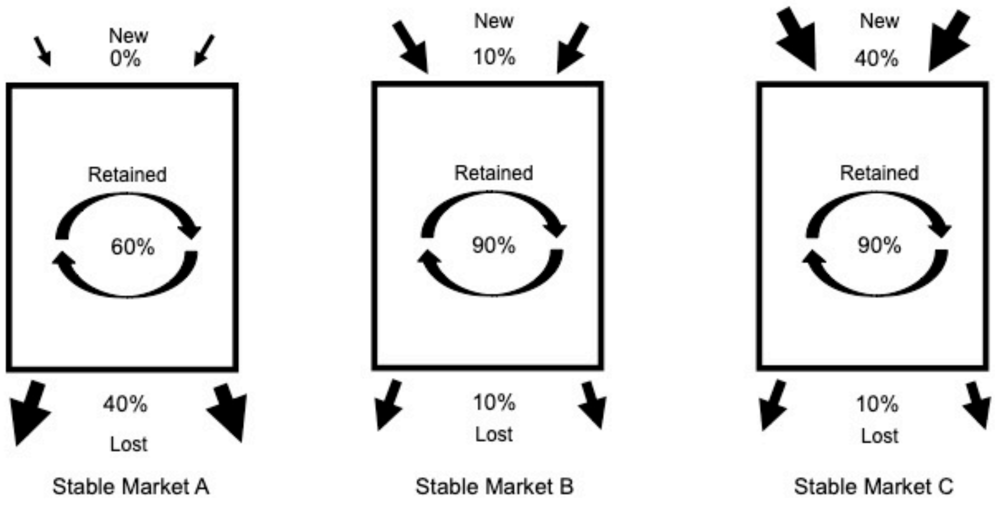

Figure 2.2 - Customer churn within a Stable market context adapted by the author from Blythe (2009)

Although seemingly stable, the level of fluctuation within each market is quite different, becoming increasingly more porous from market A to market C. Markets experiencing high volumes of defections will struggle to grow just as it would be difficult to fill a ‘leaky bucket’. The leaky bucket theory is also identified in the work of Ehrenberg (1972), for the way it limits the possibility of ever achieving the goal of 100% repeat-buying. The leaky bucket theory dictates that there may be a natural turnover under stable conditions where lost customers are replaced by new customers. Or in fact, perhaps lost buyers simply lapse for a period and then return into the market again.

Like customer churn, the use of the leaky bucket theory although briefly documented in marketing literature, is seldom focused on in isolation. It is addressed as part of other papers on customer lifetime value (2.2.3) and customer switching behaviour (2.2.4).

2.2.3 Customer Lifetime Value

One of the foundational theories behind market churn is the concept of customer lifetime value (CLV). This reflects the dynamic nature of a customer’s relationship with a service over time. Firms experience both short and long-term customers with variations in how profitable these customers are (Reinartz and Kumar, 2003, Kotler, 1994). Studies of customer lifetime value have been conducted in both contractual service settings (for example, Bolton et al., 2000) and non-contractual service settings (for example, Reinartz and Kumar, 2000). In non-contractual service settings, customer lifetimes are more difficult to predict as there is no obligation to the customer to repurchase (Reinartz and Kumar, 2003). The underlying principle is that the longer a customer is retained, the more profitable they are likely to become over time (Reinartz and Kumar, 2000). However, each customer will reach an individual point where they become unprofitable and will either partially or completely cease business with that provider.

A purpose behind studies of market churn is to extend the profitable life of a customer within a business for the longest possible length of time. Studies have generally focused on the perceived link between customer satisfaction and the length of time a customer is retained (Bolton, 1998, Lee et al., 2001). However, many factors might influence the decision of individual customers to defect from a business or service, both exogenous (those external to the service provider) and endogenous (internal factors that can be influenced). Of equal importance, though given limited attention within the existing literature, are the reasons that a customer might commence with a business. This is often associated with leaving a competing business, and causes customers to weigh up the costs of making a change, or 'switching costs'.

2.2.4 Customer Switching Behaviour and Switching Costs

Customer switching behaviour is the phenomenon of customers that dramatically change their behaviour by leaving or commencing use with a company. Switching behaviour is measured by understanding the 'switching costs', or the difficulty that someone would experience in defecting and then commencing with a new service to meet the same need. Switching costs can be characterised by both transaction costs (sign up fees, time spent filling out paperwork) and search cost (seeking out information on alternatives). The availability of information in current times has made customers more transient than ever, as they can re-evaluate their choices with little cost or effort, resulting in all markets becoming more porous (Tamaddoni et al., 2014, Holtrop et al., 2016). Services with high switching costs can make people less likely to switch and can result in a false impression of customer loyalty (Lee et al., 2001). Non-contractual services might have low switching costs and create the impression that customers are highly prone to fluctuate. The time periods used to calculate customer lifetimes must be carefully considered to reflect the average length of time between repeat purchases to ensure measurement does not under or overestimate occurrences of acquisitions and defections.

It is also important to note that in non-contractual settings, it is difficult to define the changes within a customer base. Using the prescription drug market as an example, doctors prescribe medications to buyers from a limited selection of brands within their armamentarium (a collection of resources available to medical professionals). This creates a divided loyalty where it is difficult to discern where customers have made an active switch to an alternative market, or where a doctor has simply shuffled their armamentarium (Riebe et al., 2014, Sharp et al., 2002). This introduces the concept of 'repertoire' markets which have few solely loyal buyers with most buyers spreading their requirements across several brands (Sharp et al., 2002). This differs from 'subscription markets' where churn has been most commonly applied as there is a high level of loyalty and most buyers will use only one brand for all their requirements. Riebe et al. (2014) state that 'distinguishing between a change in the underlying choice propensities and the display of polygamous loyalty is not possible'. Given these background elements, we can begin to understand the strength and limitations of customer churn as a concept for measuring customer markets.

2.2.5 Strengths and Limitations of Customer Churn for Application to Public Transport Markets

This section provides a review of both the strengths and the weaknesses which need to be considered in seeking to apply the concept of customer churn to public transport markets.

One of the greatest strengths of measuring churn is understanding the natural level of change in the market. This allows for benchmark rates of acquisition and defections to be identified within an industry. Capturing this information allows for the quick identification of unusual levels of market change (Riebe et al., 2014). This can also indicate when service-based issues may need to be addressed to stem customer defections. This may benefit public transport operators whose markets are impacted by both exogenous and endogenous factors in differentiating impacts and creating targeted strategies to grow markets.

There has been some success in attempts to apply churn in non-contractual settings such as retail markets (Buckinx and Van den Poel, 2005), pre-paid telecommunications services (Tamaddoni et al., 2010) and public transport (Mason et al., 2011). Once the target rate of churn is set, operators can then identify variables for the estimation of churn prediction, including information on customer history and lifetime predictions (Ballings and Van den Poel, 2012, Glady et al., 2009), attitude/satisfaction (Bellman et al., 2010) and switching costs (Wieringa and Verhoef, 2007). This can then be used to target customers for improved retention rates and builds on the existing knowledge of public transport market operations as discussed further in section 2.3.

There are limitations associated with the application of customer churn. Brady (2014), identifies that one of the key issues with measuring churn is identifying the relevant variables for use - in particular temporal elements. Another limitation is that only limited information is provided by focusing on a single relationship between lost and retained customers (Kamakura et al., 2005, Tamaddoni et al., 2014). Despite only measuring a single relationship, the temporal scale for the measurement of customer churn remains difficult to identify as it changes between industries and settings (contractual vs non-contractual). Tamaddoni et al (2016) comment that the current way churn is defined implies that the loss of customers is a permanent state. This makes it less relevant to non-contractual service settings, such as public transport, where customers have no obligation to repurchase or re-patronise a service within a certain time frame.

This assessment of the strengths and limitations of customer churn identifies that there is still value to be gained from its use in marketing. However, adjustments would be beneficial to create a more reliable and holistic measure suitable for public transport markets. These adjustments will be discussed in Chapter 3: Research Approach, where a new concept for measuring market change is developed.

2.2.6 Key Conclusions Drawn from the Marketing Literature

These findings have identified several considerations for the application of customer churn to public transport markets. Public Transport Markets are largely a non-contractual setting, where riders are not required to enter a contract to use the service, though some may choose to (for example, registered yearly smart cards). The public transport market is closely related to a “repertoire market” where there are multiple options available to use simultaneously (Sharp et al., 2002). As such, the measurement of customer churn must focus on a dynamic approach that reflects the level of variation expected in the market.

Public transport has multiple modes of transport which they are in indirect competition with each other. This influences the considerations of switching costs, as they are significantly reduced as individuals can move between multiple modes where possible or required. As discussed by Sharp et al. (2002) with regards to the pharmaceutical industry, polygamous markets, where users can spread their patronage across multiple sources, makes identifying customers that are gained or lost very difficult. This may be linked to the limited number of studies measuring churn within public transport markets, which will be discussed in section 2.3.

2.3 Public Transport Market Perspectives

Understanding ridership is fundamental to understanding public transport markets, and patronage goals generally seek to maximise patronage of all types (Walker, 2008). Despite the attention given to attracting new riders, market churn has had limited application to public transport markets in the academic literature. This section explores the existing studies measuring churn in public transport and market segmentation approaches. As well as these core areas, this section also reviews fields related to the understanding of measuring market change, including the characteristics of public transport markets, ridership loyalty and seasonality.

2.3.1 Existing Churn Measurement in Public Transport

There have been limited studies that have applied the concept of customer churn to public transport markets. The most notable study has been completed by Mason et al. (2011). This industry report reviews the occurrence of churn within the British Rail commuter and leisure markets. They identified the limitations of public transport markets typically measuring the net change in markets rather than the variations across individual users. Unlike churn studies that only consider defections (lost users) this study also accounted for changes in the demand of both new customers and retained customers. As such, this approach measures changes in demand within British rail markets rather than just measuring the proportion of lost customers.

The method for obtaining data included a profile survey to identify rail customers over the last two years. This survey was then used to create a sample for a more detailed panel survey to detect the influence of churn and any reasons provided. The study of commuter travel identified the percentage of customers that were ‘new' (market entry), ‘lapsed' (market exit) and ‘loyal' (retained) customers. This analysis divided users into commuter and leisure travel and collected data for each group separately. Surveys followed users over two years, between February 2009 – February 2011.

For the commuter market. The following definitions were used for retention, market entry (new users) and market exit (lost users):

1) “Customer retention (proportion of customers who were ‘loyal’ to rail commuting - were in the 2009 and 2011 customer base)
2) Market entry (proportion of customers who were not commuting by rail in 2009 but are commuting by rail in 2011)
3) Market exit (proportion of customers who were not commuting by rail in 2011 but were commuting by rail in 2009).” (Mason et al., 2011, p.g. 4)

Although not clear in the above, customer retention within this study was calculated as the number of new users and the number of loyal users.

The results of this commuter study identified the following rates of customer churn, as shown in Table 2.1.

Table 2.1 - Results of Customer Churn in British Rail Commuter Markets 2009 – 2011 (Mason et al., 2011)

	Two Year Churn Measures		
	Customer Retention	**Market Entry (New Users)**	**Market Exit (Lost Users)**
Total	77%	23%	23%
London	83%	17%	15%
South East	79%	22%	21%
Rest of UK	74%	26%	26%

These results identify the following:

- The proportion of users that entered or exited were roughly equal in all locations surveyed
- The percentage of loyal users and customer retention represented most users, at over half of the population for all modes and up to 80% of commuters in London.

Unfortunately, for the sake of comparison and understanding the broader patterns within rail markets, less accurate data was collected for the proportion of leisure users affected by churn. The leisure travel section focused on identifying whether users were travelling more, less or about the same number of trips as the previous year. This identified that 58% of people made about the same number of trips, 25% made more trips and 17% made fewer trips.

This research of Mason et al (2011) provided findings across two years of survey data between 2009 and 2011. However, due to the changing external conditions within these time frames, the underlying level of churn may be over-estimated and this was a noted limitation of the study. This sensitivity to external conditions emphasises the value of regularly measuring and reporting on churn to detect influences and identify the underlying rate of market churn at a finer level of detail rather than relying on a less frequent measurement.

There have been no other studies that have applied churn to public transport markets as directly as Mason et al, although it has been applied to other related areas. There have been studies by both Saleh and Farrell (2007) and Chatterjee (2011) on 'asymmetric churn' in the public transport industry. Asymmetric churn is used to measure travel behaviour as an unequal two-way process that changes over time. This is most frequently used to capture changes in travel mode but can also refer to changes in departure time or other travel elements (Saleh and Farrell, 2007). Asymmetric refers to the fact that changes will seldom be in equilibrium, resulting in a net change that is positive for certain mode/s but negative for other modes. The aim is to witness asymmetric churn in favour of public transport to improve sustainability. This accounts for the fact that public transport is not just competing with itself but also the private car and active transport use.

Studies of asymmetric churn question whether variations in travel are evidence of asymmetric churn or simply usual travel behaviour, which naturally varies. As an example, it is difficult to capture or categorise a car user that chooses to catch the bus while they are unable to drive or their car is getting repaired as they have made a temporary rather than permeant shift. Currently, categorisation of users depends largely on your choice of definitions and the rules put in place around measuring asymmetric churn. Chatterjee (2011) identified that changing circumstances can be further broken down as a product of history (circumstances and behaviour) and persistent individual differences (unobserved heterogeneity).

Ortega-Tong (2013) completed a thesis on customer churn of London Oyster Cards (smart cards), that treated smart cards as individual users. This study defined churn as the attrition of oyster cards over a measurement period and studied the variations between registered and unregistered smart cards. They identified several important temporal variables including travel frequency and journey start time, spatial variability (origin frequency and travel distance), activity pattern variability (activity duration) and sociodemographic characteristics (special discount cards) (Ortega-Tong, 2013, p.g. 145).

This work found a significant drop in the number of active Oyster Cards observed between consecutive months, which was largely attributed to occasional user cards that exit the system at a faster rate than cards of regular users who retain their cards for longer.

Of the available studies, all found evidence of churn occurring within public transport markets. No existing studies looked at individuals whose ridership patterns didn't fit into the categories of new, lost or retained users, for example users that start, stop and then return to travel repeatedly. This may result in dramatically over-estimating the impacts of churn, as was a noted limitation of Mason et al. (2011) by counting natural variances in ridership as churn instead of an ongoing process of change. It is suggested that further refinement to the concept and measurement of customer churn is required for its appropriate application within the public transport industry. To guide this, this literature review identifies studies that seek to segment public transport markets based on differing sets of conditions. These approaches and what we can learn from them are reviewed in the following section.

2.3.2 Similar Ridership Market Segmentation Approaches

As in marketing, the analysis of public transport ridership involves clustering users into many different types of market segments, although rarely through the lens of customer churn. The objective of market segmentation approaches often being to reduce the study data into a manageable number of mutually exclusive groups with similar and clearly defined characteristics (Anable, 2005).

The most obvious classification is separating the population into public transport users and non-users. However, there is limited research attempting to understand the preferences and choices of the non-user population, despite the potential for this group to become new public transport users (Krizek and El-Geneidy, 2007). Transit users can then be further broken down into frequent and non-frequent users, however, this is most commonly reviewed as sub-categories after dividing riders into choice and captive ridership (Krizek and El-Geneidy, 2007, Jacques et al., 2013, Jin et al., 2005).

At its simplest, choice riders are those that are considered to have numerous transport options available to choose from for their intended travel purpose. Captive riders only have the option of catching public transport (or even a specific mode of public transport) for their trip. Choice riders are understood as a diverse socio-demographic group, where captive riders are generally thought of as the low income, elderly, children, differently-abled or those that do not own a motor vehicle (Krizek and El-Geneidy, 2007).

The application of the captive-choice segments to real-world public transport markets has received significant criticism. The framework leads transport operators to assume that rider defections must come from the choice rider population, as they are more sensitive to changes in service quality and cost than captive riders (Beimborn et al., 2003, Jacques et al., 2013). Transport operators can also fall into the trap of assuming they will always retain a captive rider population (Beimborn et al., 2003). Further there is the assumption that any defections from captive riders are from them moving into the choice rider category, for example by acquiring a motor vehicle (Krizek and El-Geneidy, 2007).

Krizek and El-Geneidy (2007) sought to improve the concept by developing alternative categories of users: irregular and regular captive users and irregular and regular choice users. Non-users were classified as irregular and regular potential transit users, and irregular and regular captive auto users. Similarly, Jacque et al. (2013) used a web survey to develop four new segments of market use;

Captivity (true): Low level of trip practicality and low level of preference for this trip

Utilitarianism: High level of trip practicality and low level of preference

Dedication: Low level of trip practicality but a high level of trip preference/satisfaction

Convenience: High level of practicality and a high level of trip preference.

This approach is beneficial in applying a continuum of travel segment to an individual's current trip, as opposed to attempting to understand all trips.

Beyond the measurement of captive and choice riders, there is some evidence of customer churn in segmentation approaches. For example, Saleh and Farrell (2007), noted there were four categories of people that would occupy a seat: loyal passengers, new passengers, lost passengers and an unoccupied seat. These categories are illustrated as to how they look on a bus service in Figure 2.3. Each of these categories should be treated separately in terms of both analysis and strategies to influence customer retention.

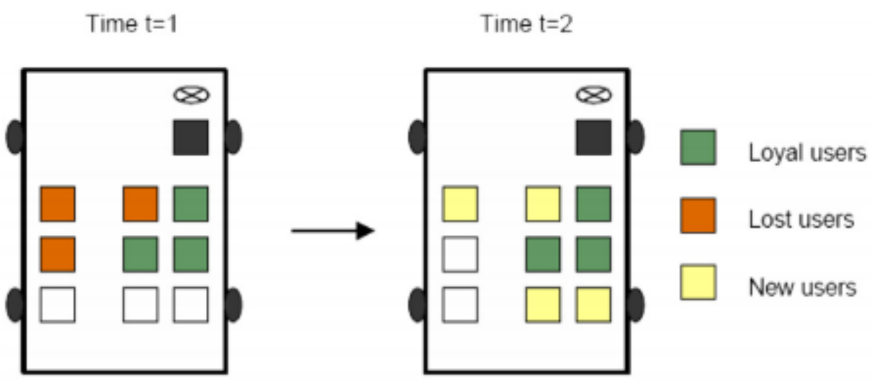

Figure 2.3 Diagram of asymmetric churn categories on a bus service (Chatterjee, 2001)

Figure 2.3 highlights how segmenting public transport passengers using customer churn is impacted by the element of time, as the type of users on the bus are quite different at the two points in time where measurements are taken. What is not known is why these users are changing over time, and whether users are appearing repeatedly within different segments. As an example, a new user, might become a retained user for a period before later being lost. This approach identifies the potential for investigating churn based analysis of market change with opportunities to refine the clarity and depth of information gained from such approaches.

Finally, Anable (2005) proposes an attitude theory based approach to segmenting users, where users are segmented based on their attitudes and motivations to use various modes of transport. In this approach users are surveyed based on attitudinal and then placed into segments through a post hoc approach, where some form of multivariate statistical analysis is used to identify segments (e.g. cluster analysis). This offers an alternative to the choice/captive rider approach, where users were instead segmented into four types of car users and two types of non-car users based on their different responses regarding 'environmental concern, participation in pro-environmental behaviour and moral obligation' (Anable, 2005). This offers some benefits when compared to traditional a priori segmentation as it does not assume to false assumptions of homogeneity within the pre-selected groups and can remove the bias caused by this when interpreting behavioural tendencies for example non-car users could be split into those that were 'carless crusaders' (e.g. by choice for environmental reasons) and those that were 'reluctant riders' who were required to take public transport for health or financial reasons. Though this is beneficial for the explanation of transport user behaviour, this approach without some a priori segmentation is less useful for measuring changes in public transport ridership.

As there are a limited number of similar market segmentation approaches, it is also important to review other areas that measure changes in ridership and travel patterns. As such, the next sections explore habitual and non-habitual travel patterns, loyalty and seasonality and how they influence public transport markets as well as ridership decisions.

2.3.3 Habitual and Non-Habitual Travel Patterns

Understanding travel patterns is important for both operators and policy makers as it can assist with growing ridership markets through identifying areas to target for inciting a behaviour change (Bamberg et al., 2003, Chen et al., 2011a). Travel is regularly thought of as habitual behaviour, that continues until a prompt emerges that requires re-evaluation. However, each time a person is required to travel, they mentally conduct a cost-benefit analysis of the most appropriate journey to take given their major concerns, for example considering the impact of cost, convenience and reliability for the modes they have access to (Guiver, 2007). Thøgersen (2006) focuses on the development of travel habits, arguing that the cost-benefit analysis of a travel behaviour becomes habitual to the individual after around a month of repetition. This has interesting implications as it can result in a 'tunnel vision' approach to travel decisions, where a habitual car user may not be aware that there is a more efficient and direct public transport route available to them as they no longer seek out this information as the decision has become automated (Walker et al., 2015, Verplanken et al., 1997, Van Exel and Rietveld, 2009)

Disruptions to the context in which the routine occurs, particularly through significant life events, have been shown to significantly weaken habitual travel behaviour as they change the environmental cues that trigger reflexive behaviour (Walker et al., 2015). Additionally, the introduction of new information is likely to change travel decisions even in the case of habitual behaviour (Bamberg et al., 2003). This is an under-represented impact of real-time data available for public transport users as it provides greater ease, convenience and flexibility to how travel decisions are made, particularly with regards to mode and route choice (Buehler et al., 2017). We can hypothesise that this availability of real-time information may have increased the rate of customer churn within public transport markets. Behrens and Mistro (2010) also identify the difficulties in separating non-permanent 'churn' (A temporary pause in travel for holidays or illness) from deliberate 'churn' events (an active decision to stop using a public transport mode). This makes it difficult to truly understand the extent to which churn impacts on public transport markets.

2.3.4 Understanding Customer Loyalty in Public Transport

Studies of public transport ridership show that a small number of users who travel at high frequencies often represent a high proportion of trips (Bagchi and White, 2005). In the context of public transport, 'loyalty' refers to a commitment to re-patronise a preferred service consistently in the future (Tao et al., 2017a, Oliver, 1999). Loyalty is still a relatively small interest area for studies of public transport. Much of the focus in public transport is predicting loyalty using satisfaction with services (St-Louis et al., 2014, Olsson et al., 2013, De Vos et al., 2016, Abou-Zeid and Fujii, 2016), behavioural intentions (Lai and Chen, 2011, Tao et al., 2017a) or both (van Lierop and El-Geneidy, 2016).

The Dick and Basu (1994b) model of loyalty looks at loyalty as the relationship between a user's relative attitude and repeat patronage. The relationship between attitude and patronage can then be used to create four unique ridership segments as illustrated in

Figure 2.4.

		Repeat Patronage	
		High	**Low**
Relative Attitude	**High**	Loyalty	Latent Loyalty
	Low	Spurious Loyalty	No Loyalty

Figure 2.4 – Reproduced Diagram of Dick and Basu Four Segment Loyalty Model (Dick and Basu, 1994b)

Allen (2004) takes a slightly different approach to understanding loyalty in an operational context. This approach includes the relationship between the demonstrated repurchase or reuse of a product/service, the likelihood of recommending to others, and overall satisfaction (Allen, 2004). The focus of many studies of loyalty in public transport has been on the link between satisfaction with the service and loyalty. Few, if any, studies have focused on loyalty in terms of witnessed repeat patronage. Webb (2010) identified that this is likely due to the historical limitations in being able to track individual rider's public transport use.

Measurement approaches to understanding loyalty in response to external or internal changes are of the greatest interest to understanding customer churn. The connection between satisfaction (measured through variables such as safety, comfort or cleanliness) and loyalty and how that varies for captive, choice and 'captive by choice' transit users was measured by Van Lierop and El-Geneidy (2016). The results of this study identified that service quality improvements would influence the loyalty of the groups (choice, captive or captive by choice) in different ways. However, changes targeted at one group such as captive riders may also positively influence ridership in other groups. Similar studies have found that loyalty was positively influenced by good experiences of service and that loyal passengers were less likely to shift to alternate modes (Tao et al., 2017a). Other studies have measured the impact of external changes (Hoang-Tung et al., 2014) or the likelihood of users migrating to a new mode (Bass et al., 2011). The increasing availability of smart card data has also increased the ability to measure and predict changes in loyalty over time. As an example, Trépanier et al (2012) found through hazard modelling that the rate of terminating a smart card increases over time.

Although there is existing research literature surrounding customer loyalty in public transport, there is still a limited understanding of many aspects of loyalty. In particular, there is a reliance on cross-sectional surveys which limits the ability to observe real-time changes, and a lack of individual data to appropriately segment markets (Tao et al., 2017a). Limitations also arise from the use of aggregated assessment, as it is increasingly difficult to capture individual decisions around preferred mode choice (Bass et al., 2011). A final concern is the lack of non-public transport users in existing panel studies, limiting our ability to understand how these users behave or how they may be targeted in marketing programs (Transit Cooperative Research Program, 1999).

The focus of most papers on public transport loyalty is on identifying the service based influences of loyalty rather than measuring loyalty itself in terms of actual ridership. The nature of loyalty of public transport users is yet to be fully understood concerning its relationship to passengers' actual and intended use of public transport (Tao et al., 2017b). Tao et al (2017) question whether we can identify if loyalty is a product of preferred or constrained choice. This reflects many of the critiques of labelling users as either captive or choice riders as it cannot be assumed that their situation will not change.

Another constraint to consider when measuring the loyalty of user's public transport ridership is the impact of seasonality on public transport markets. This is discussed in the following section.

2.3.5 Seasonality in Public Transport Markets

One of the key questions when measuring individual ridership over time is differentiating between active changes in travel decisions and natural variations in travel that occur over time. This leads us to the concept of seasonality which seeks to measure variations in transport ridership as they can be linked to different temporal factors or seasons (Bocker et al., 2013). As noted by Briand et al. (2017), seasonality goes hand in hand with mobility. As such, studies that measure the changes in travel patterns over the course of the year will by nature be impacted by variations in seasonal use of public transport. This is based on the understanding that transport demand is constantly changing depending on days of a week, and different times of the year (Amiripour et al., 2014, Ahmad Termida et al., 2018). Seasonal changes capture changes in demand relating to holiday periods as well as weather that is inclement for travel (too hot or too cold). Unlike the other attributes that impact on travel demand, such as changes in fare/ticket price or service levels, seasonality is often neglected in the literature on public transport. Alternatively, seasonality is a common consideration for other transport industries such as aviation where there are well known high demand periods of travel in the summer months, around school holidays, the festive seasons, on weekends and around large sporting events (Merkert and Webber, 2018). This is likely a greater focus of study in aviation than public transport as though both industries have fixed capacities, airlines are able to respond to seasonality with dynamic pricing (Merkert and Webber, 2018). Overall, the study of seasonality opposes the argument that travel demand is randomly varied over the course of days, weeks and months and instead reflects broader context based changes.

Seasonality can be measured at a range of different time periods with varied results. The term seasonality implies a focus on calendar seasons or at the least a period of consecutive months that showcase a similar travel characteristic. For example in the northern hemisphere when looking at regular weekday users (commuters), there are clear seasonal variations during summer and over the Christmas period (Chu, 2015). However, while looking at weekend travel more generally it is consistently around 10% of total use for the entire year with occasional peaks often directly explained by sports or other entertainment events (Chu, 2015).

An important aspect of seasonality studies is also the activity being undertaken. Tang and Takhuriah (2012) analysed the impact of seasonality on human activity-travel behaviour on Chicago's public transport. They found evidence that bus ridership was higher in autumn (September-November) and spring (March-May), and lower during summer and winter months (except for February). However, we are aware that some modes such as bus, which rely heavily on school travel, are greater impacted by seasonality effects on ridership. Regardless of school travel, the bus ridership findings show the close connections between seasonality impacts and weather impacts (Ahmad Termida et al., 2018). Other studies have found limited variation in daily travel behaviour in the Netherlands, particularly for routine trips (e.g. work) (Kitamura and Hoorn, 1987). Studies have identified greater variation due to seasonality when studying recreational activities (weekends) (Bhat and Srinivasan, 2005). As weekend trips are responsible for a small proportion of public transport travel, macro studies of public transport use are not overly affected by small, individual changes in travel due to seasonality effects where weekends are the most impacted (Chu, 2015).

There is also evidence of regular day-to-day variability in travel behaviour due to riders that have different needs and commitments on different days (Susilo and Axhausen, 2014). The measurement approach in studies of seasonality must carefully consider the timeframes used so that multiple weeks are covered, allowing for the capture of routine daily behaviour (Schlich and Axhausen, 2003). Schlich and Axhausen (2003) found in a Dutch panel that shopping had identical daily patterns on 5 or more days in a two week study period, measured 6 months apart. This implies that seasonality may not have as substantial an impact on overall ridership as once thought.

Seasonality focuses on the intricacies of existing public transport user behaviour, changing day-to-day or week-to-week in response to seasonal impacts. It is predominantly focused on existing users with a limited scope to measure new and lost users. Customer churn is a broader approach focused on the market, which will see some influence of seasonality. It is suggested that different measurement periods are reviewed, similar to Schlich and Axhausen (2003) to account for seasonality differences and variation in travel.

2.3.6 Key Conclusions from Public Transport Studies

There have been a small number of studies that can be connected to the marketing concept of customer churn within the field of public transport, yet there is still significant scope for further investigation of the concept. The application of churn should learn from both existing direct attempts to apply churn (Mason et al., 2011) as well as broader market segmentation approaches to capture the dynamic nature of public transport ridership. Importantly, measures should go beyond the traditional churn approach of focusing on the relationship between retained and lost customers but also account for new riders that are acquired as well as the natural level of internal variability. Understanding internal variability is important as it can reduce the impact of incorrectly separating non-permanent churn from deliberate churn events (Behrens and Mistro, 2010).

Based on the characteristics of public transport markets four key segments should be considered: new users, lost users, retained users and returning users (internal variability). Considering all four of these elements would allow for the measurement of market change, rather than simplified measures of churn which focus on the relationship between lost and retained users only.

Table 2.2 provides a synthesis of studies that are considered to most directly relate to market churn against the four identified categories of market change. Studies were selected that could be most closely linked to the study of customer churn and must involve the measurement of individual patterns of patronage of a service.

Table 2.2 - Synthesis of Literature against Four Key Elements of Market Change

Reference	Acquisitions	Defections	Retention	Internal Variability	Public Transport Markets
MARKETING/CHURN					
(TAMADDONI ET AL., 2016)	No	Yes	Yes	No	No
(TAMADDONI ET AL., 2014)	No	Yes	Yes	No	No
(ATHANASSOPOULOS, 2000)	No	Yes	Yes	No	No
(REINARTZ ET AL., 2005)	Yes	No*	Yes	No	No
(REICHHELD AND SASSER, 1990)	No	Yes	Yes	No	No
(RIEBE ET AL., 2014)	Yes	Yes	Yes	No	No
PT/CHURN					
(MASON ET AL., 2011)	Yes	Yes	Yes	Yes	Yes
(SALEH AND FARRELL, 2007)	No	No	No	Yes	Yes
PT/ NET MARKET CHANGE					
(CURRIE AND WALLIS, 2008)	Yes	Yes	No	No	Yes – Bus market focus
(CHEN ET AL., 2011B)	No	No	No	Yes – Aggregate, macro-level	Yes
PT/ LOYALTY					
(TRÉPANIER ET AL., 2012)	No*	No*	Yes	Yes	Yes
(VAN LIEROP AND EL-GENEIDY, 2016)	No	No	Yes	No	Yes
(BASS ET AL., 2011)	No	Yes	Yes	No	Yes
(TAO ET AL., 2017A)	No	No	Yes	Yes	Yes
PT/ VARIABILITY IN BEHAVIOUR					
(CSIKOS AND CURRIE, 2008)	No	No	No	Yes	Yes
(BRIAND ET AL., 2017)	No*	No*	No	Yes	Yes
(CHU, 2015)	No	No	No	Yes	Yes

*Although measured, not utilised in discussion.

This synthesis shows that except for a single study by Mason et al (2011) there is a gap in the academic literature of studies that measure the interplay between all four components of market change within the field of public transport. Addressing this gap by developing an updated tool to measure the market change in public transport markets could provide operators with a beneficial measurement tool to inform decision making. As such, it is important to review existing approaches to measuring public transport ridership as part of this exploratory literature review.

2.4 Measuring Public Transport Ridership

There has been a large amount of research devoted to understanding and predicting travel demand (Saleh and Farrell, 2007). Within this research, there is a broad selection of measurement approaches. Most commonly, the measurement approaches fit into three categories, cross-sectional surveys (e.g. Clayton et al., 2016, Kitamura and Hoorn, 1987), longitudinal surveys (e.g. Clark et al., 2014, Csikos and Currie, 2008), and smart card data (e.g. Chu, 2015, Ma et al., 2013, Zhang et al., 2018). For this discussion, survey approaches are summarised collectively.

An issue with the existing approaches is that there are limited consistencies with the definitions for regular ridership and temporal variables between different studies. In this instance, the focus has been on approaches that measure regular travel patterns of individuals. Zhong et al. (2016) define regularity or regular travel, as a “uniform pattern, principle, arrangement, or order that repeats itself”. This definition is key as these patterns are reproducible, and thus can be used to predict ongoing transport behaviour a valuable piece of knowledge for public transport operators and transportation planning.

2.4.1 Survey Approaches (Cross-Sectional and Longitudinal)

Surveys of public transport users have been extensively used within the measurement of public transport studies. This work focuses on cross-sectional surveys (those that are completed by a single respondent at a single point in time) and longitudinal surveys (those that follow a panel of people over a period of time) (Rindfleisch et al., 2008). Both approaches have strengths and limitations which will be discussed within a public transport context.

Cross-sectional surveys are valuable as a quick and effective tool to understand what is happening within public transport markets at a point in time, often collected over a short, dedicated period. In public transport studies, they have been extensively employed to capture the values of users through surveys measuring attitudinal values and stated preferences (Clayton et al., 2016, Chalak et al., 2016, Schmitt et al., 2013). There have been some instances where cross-sectional surveys have been used to capture changes in travel behaviour, however, these have often been limited in application to measuring a response to a specific event. For example, Jacques et al. (2013) utilised a university access survey that allowed users to change their selected mode of travel to capture seasonality effects.

Traditional cross-sectional survey approaches fail to capture the temporal element of behaviour change (Saleh and Farrell, 2007, Behrens and Mistro, 2010). This can be partially overcome with approaches that rely on repeated cross-sectional waves surveys over an extended period as utilised and thus sit between a true cross-sectional and a longitudinal approach (Kitamura and Hoorn, 1987). There have also been studies that rely on participants to recall information over an extended period of up to several years (Schmitt et al., 2013, Beige and Axhausen, 2008). Beige et al. (2008) identified the power of retrospective surveys, where participants are asked to recall key behaviours, to address the time and expense limitations of longitudinal panel data. A similar study by Schmitt et al. (2015) was developed to address the issues with participant recall by allowing for multiple surveys (before and after) during a cross-sectional approach to improve accuracy. Axhausen et al. (2007) found that 1-day travel diaries overestimated the stability of the behaviour under study when compared to multiple-day surveys or observations which better account for personal variability in behaviour.

Finally, some consideration has been given to the use of longitudinal survey data in measuring passenger travel behaviour. Longitudinal data, addresses many of the limitations of cross-sectional or retrospective survey approaches by allowing for multiple measurements to be taken over time and reducing the risk of error from using a single data source (common method variance). Longitudinal data also allows for a greater level of confidence in the inferences that can be made from the research findings (Rindfleisch et al., 2008). Numerous studies have relied on extensive existing data sets such as household surveys to investigate changes in behaviour over time (Bamberg et al., 2003, Briand et al., 2017, Clark et al., 2014). However, the collection of project specific longitudinal data is prohibitively expensive and time consuming for much academic research (Rindfleisch et al., 2008). Further, longitudinal data collection requires significantly larger sample sizes, due to the increased rate of participant attrition over time (Rindfleisch et al., 2008).

Rindfleisch et al. (2008) found that, in certain circumstances, that the validity of cross-sectional data was comparable to that of longitudinal data. Further, longitudinal panel surveys that seek to answer complex questions can be impacted by low response rates as users become fatigued by poorly managed or conducted travel diary data collection (Axhausen et al., 2007). Some studies however have managed this by repeating cross-sectional studies to build longitudinal data (Tang and Thakuriah, 2012). This might suggest that a similar approach is the most practicable for meeting immediate research needs while allowing for the benefit of longitudinal data if research progresses.

2.4.2 Smart Card Data

For more than two decades, smart card data has offered a world of possibility for public transport researchers with its ability to collect huge volumes of detailed travel data. Although it offers potential, the use of smart card data has not come without limitations. This section will address what smart card data is, the potential to capture public transport user information, the measurement approaches used (including regular assumptions made) and the limitations of smart card data.

Smart Cards have become the predominant ticketing system for public transport in much of the world (Pelletier et al., 2011). As smart cards offer the ability for data to be both stored and collected, their value to public transport operators moves significantly beyond recording trips to collect trip fare/revenue (Pelletier et al., 2011). Although there are a few different types, smart cards are usually a plastic card that has an embedded microchip for functionality. This significantly improves on the durability of previous paper or cardboard ticket options and means that riders can hold and re-use smart cards for many years. This affords public transport operators the ability to record historical travel pattern data for a user that continues to hold their card. This data includes where they travel, when they travel, how often and by what mode (Dempsey, 2015). Due to personal privacy concerns, there are limited instances where additional demographic information such as the name, age, gender, address or living location is recorded. Where this information is recorded, access is restricted from the public (Dempsey, 2015). However, numerous methods have been investigated to derive or re-connect socio-demographic data to anonymised smart card data sets.

Due to the extent of the data sets collected with smart card data, there are also errors within the data itself. A common error is the desynchronization between the on-board smart card reader and the planned routes resulting in incomplete data (Pelletier et al., 2011). Another limitation is that in many cases, lost or stolen smart cards do not retain a continuous ID for tracking purposes (Briand et al., 2017). Adding to this, in many systems there is a regular systematic requirement for card replacement. This results in samples growing increasingly smaller as the analysis time frames grow larger (Chu, 2015). Smart card analysis also does not allow us to understand the motivations behind individual users travel decisions due to the current limitations of data collected (Briand et al., 2017).

Further, the value of smart card data, in terms of its completeness as a data set, varies dependant on the system that it operates within. For Melbourne and Santiago (Devillaine et al., 2012) there are completeness issues as their systems do not consistently require users to touch off when alighting a public transport vehicle for fare collection. This means that researchers need to go back and estimate alighting locations based on the data available, which has become a significant area of study as reviewed by Li et al. (2018a).

The uses of smart card data have extended well beyond a means of collecting payment, with a huge volume of studies amassed since its introduction. The applications for smart card analysis have been broad and varied including the measurement of spatial travel patterns (Ľudmila Jánošíková, 2014, Kieu et al., 2015a, Zhao et al., 2017), temporal changes in passenger behaviour (Morency et al., 2006, Morency et al., 2007, Trépanier et al., 2012, Ma et al., 2013, Chu, 2015, Goulet Langlois et al., 2016, Briand et al., 2017, Kim et al., 2017, Zhao et al., 2017), passenger group segmentation (Briand et al., 2015, Kieu et al., 2015b, Ortega-Tong, 2013), ridership prediction (Oort et al., 2015, Zhao et al., 2018) and inferring trip purpose (Alsger et al., 2018). This project builds on the findings and approaches of a broad range of similar smart card studies, with an emphasis on the use of smart card data as a means of understanding and measuring customer churn and any revised concepts for application to public transport markets.

2.4.3 Smart Card Measurement Scales

Identifying an appropriate temporal scale for analysis is a key question when smart card data to identify patterns of use, due to the immense scale of smart card data collected. Unfortunately, these are questions that are still without a consistent answer, as the scale will depend largely on the specific question being studied and the data available. With regards to the temporal scale, the approach has significantly varied with studies using yearly, multi-year (Chu, 2015, Briand et al., 2017), monthly (Ma et al., 2017), weekly or multi-week (Zhong et al., 2016, Ma et al., 2013, Zhang et al., 2018) and even daily (Tao et al., 2014) data samples as their entire time frame. A study by Devillaine et al. (2012) measures daily (weekday), weekly and a 9-year data set, using data available for different cities, while Morency et al. (2007) use 277 consecutive days.

Table 2.3 provides an overview of selected studies measuring variations in individual public transport ridership with a summary of the measurement approach. Importantly each work has been reviewed in terms of its limitations in measuring customer churn as it occurs within public transport markets as a dynamic process.

Table 2.3 - Review of Travel Pattern Measurement in a Selection of Key Literature

Article	Timeframe	Analysis Segment	Purpose	Sample	Limitations for Customer Churn
Briand et al. 2017	5 years	Weekday and Weekend Day	Analysing year-to-year changes in passenger behaviour using the case of Gatineau, Canada	82,223 cards	Only focuses on a reduced data set where cards are active for the entire 5-year period
Chu, 2015	2 years	Same day, day-to-day, week, seasonal, year-to-year	Longitudinal observations to understand travel behaviour	238 145 cards	Excluded cards that were lost within the time frame
Morency et al.,2007	1 year (277 consecutive days)	Weekly travel	illustrate the potential of smart-card data to measure spatial and temporal variability of transit	7118 cards	Measures patterns of retained users. Clusters only work with heterogeneous patterns,
(Ma et al., 2017)	One Month	Day, Hour	Understanding commuting patterns using the case of Beijing	500,000 riders 37, 001 cards used to test algorithms	Segmentation into commuting patterns achieved but minimum details of each group
(Tao et al., 2014)	Single Day	Hour	Spatial- temporal dynamics	515,435 transaction records	Focus on the visualization of travel paths Unlikely to be detect churn from one day of travel data

This variation in approaches identifies that the time frame and analysis segment is selected to best match the analysis purpose and the data available. As there is no clear guidance for the temporal elements required to measure customer churn, this is a clear gap that requires further testing as a component of this study. This gap is further supported by the findings of Chikaraishi et al. (2013) who argued that it is hard to obtain an optimal travel survey design for multi-day and multi-period panels because of the relatively little data available on changes in travel behaviour to date.

As well as needing appropriate timeframes, smart card data analysis is dominated by ‘rules-based' processing, often set arbitrarily in the absence of, for example, trip activity information. There is a need for service providers to undertake verification surveys to check data inferred from smartcard analysis against actual behaviour—for example, assumptions made about transfer time between successive boarding's in the same network, or trip rates by concessionary pass holders (Bagchi and White, 2005). This is also consistent with the requirements for verification approaches in the measurement of customer churn (Ballings and Van den Poel, 2012).

2.4.4 Key Conclusions from Measurement Approaches

Although a simple concept, the literature surrounding market churn and individual travel patterns highlights how difficult it is to measure in public transport markets. Both primary data collected through longitudinal or cross-sectional surveys and the use of smart card data come with their own set of strengths and weaknesses.

Surveys can be beneficial in collecting targeted data and providing the reasons behind choices that have been made. However, there is the potential of low response rates, survey response bias or fatigue and issues with accuracy. Smart card data often heralded as the solution to issues with surveys, provide large though imperfect data sets which can capture travel patterns. The quality of smart card data is dependent on the context that the data is collected within and it cannot provide information on the reasons behind travel patterns. Therefore, the research uses a mixed-method approach to best address the limitations of different approaches and assist in identifying the simplest approach for measurement.

The measurement period is a key element to be refined, as there is no consistent approach in the literature. It appears that many studies are required to adjust their measurement approach to best reflect the data that they have available and the intended purpose. Many also utilise several different temporal units of analysis for comparison within the study (Briand et al., 2017, Chu, 2015, Ma et al., 2017). What we are aware of is that we are looking for regular travel patterns that repeat themselves, as defined by Zhong et al. (2016). Multiple measurement periods and units of temporal analysis should be explored for measuring public transport markets.

2.5 Conclusions and Implications for Further Study

There have been limited direct applications of the marketing theories of customer churn to public transport markets. A review of the literature has identified that both studies of market change and studies of public transport ridership have similar objectives, and as such the application of a churn based approach offers numerous benefits. There is ongoing contention in the marketing literature as to whether focusing on customer acquisition or customer retention is more beneficial to growing customer markets and as such, there are very few studies that seek to understand the influences of both**. To address this, the development of an updated concept of customer churn that reflects the characteristics and unique case of public transport markets is proposed. The concept**

measures the impact of new, lost, retained and returning users (internal variability) to develop a complete picture of market changes. The updated churn measurement approach should also consider the reasons why customers join, defect and re-join public transport as a continuous cycle rather than discrete actions or an ongoing state of retention.

Applying an adapted concept of customer churn can offer new insights for public transport operators and policymakers. This includes the provision of new information as currently the number of new and lost customers at any point in time is not measured in favour of identifying aggregate market growth or decline. Further research should focus on developing a simple tool for measuring the patterns of customer fluctuation that can be implemented regularly by public transport operators to check and understand market health. This will also assist in improving the clarity and refining of existing modelling of public transport ridership and may be a valuable tool for understanding the influences and impacts of market change.

Chapter 3: Framework Development and Research Approach

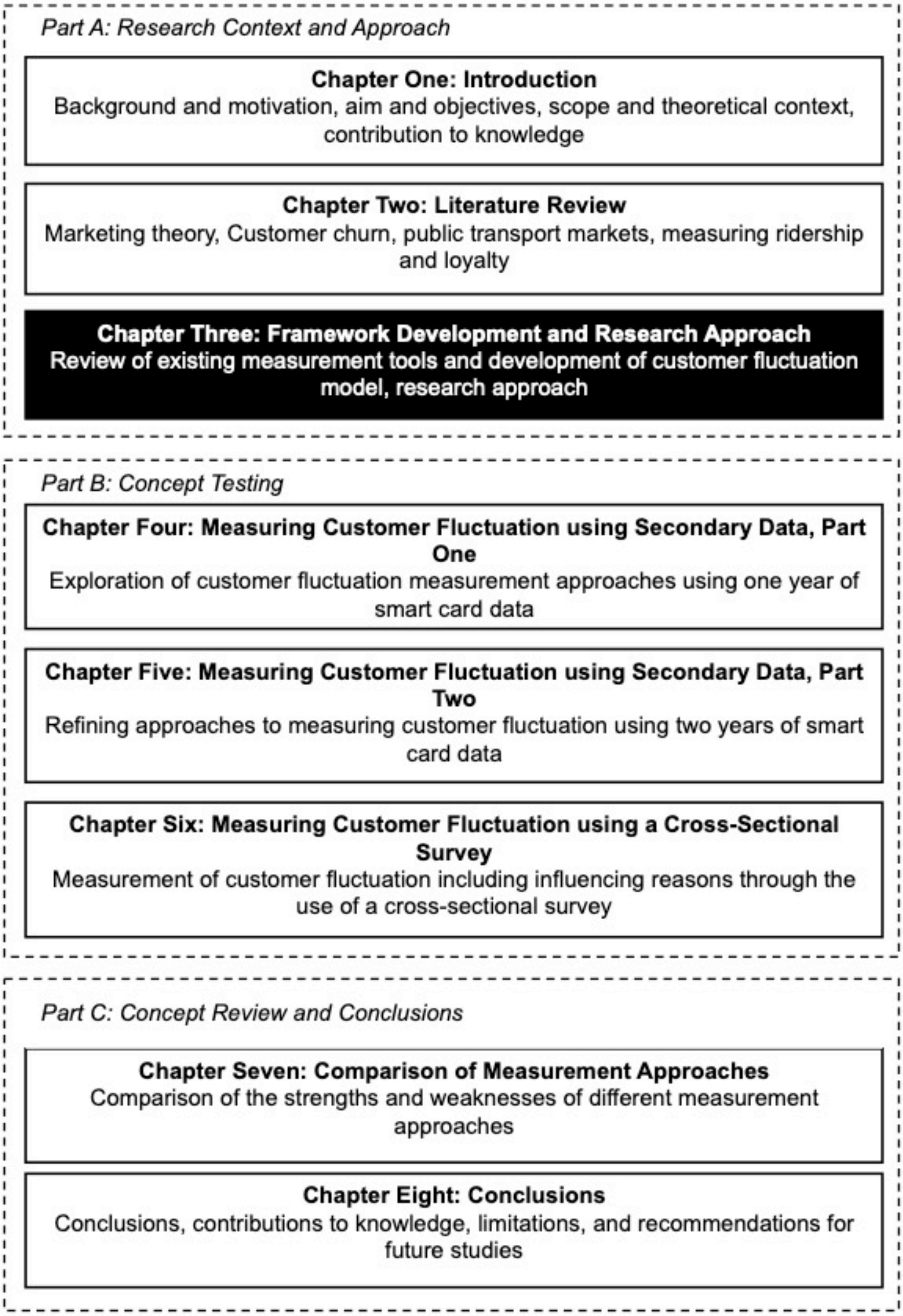

Figure 3.1 - Position of Chapter 3 in thesis structure

3.1 Introduction

Chapter 2: Literature Review provided an overview of existing research related to the measurement of customer fluctuation and customer churn. This review identified a knowledge gap in the use of approaches to measure market change that measures the four key components of customer behaviour. The review identified an ongoing contention in marketing as to whether attracting new customers or retaining existing customers is more beneficial to grow customer markets. However, this focus limits the ability to measure and understand market change. By investigating measurement approaches that identify all key behaviour change elements of public transport use, the proportion of new, lost, retained and returning customers, there is a potential to better understand market change. To do this, a new concept is proposed, termed 'customer fluctuation'.

This chapter provides a general overview of the research design that has been developed for this thesis to achieve the research aim and address the research questions. The methods that have been selected have been informed by the findings of the Literature Review. This chapter begins by restating the overarching aim of the research and associated research questions (3.2). This is followed by section 3.3, which identifies the context of this research. Section 3.4 is used to develop and define the concept of customer fluctuation. Section 3.5 then provides a rationale for why customer fluctuation should be explored as a new measurement approach. Section 3.6 identifies existing measurement tools and section 3.7 outlines the research approach used in this thesis. Brief conclusions are offered to end this chapter.

3.2 Aim of Research

There are two research aims for this thesis, which are reiterated here as follows:

I. To develop, measure and apply a new concept for market change analysis based on market change segments

II. To explore the potential of this approach as a means of improving market change analysis

To answer the research question, five key research objectives are proposed:

RO1. To understand conventional measures of market change in the fields of marketing and public transport and explore the benefits and drawbacks of these methods

RO2. To develop a new concept for market change analysis based on market change segments

RO3. To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne

RO4. To explore behavioural factors influencing market change segments using survey data for metro Melbourne

RO5. To assess the potential of the new approach for application in the industry

A detailed assessment of how each chapter addresses each research question is provided in Table 3.1

Table 3.1 - Overview of How Each Chapter Addresses the Research Aims and Objectives

Research Aims	Research Objectives	Chapter 2: Literature Review	Chapter 3: Framework Development and Research Approach	Chapter 4: Measuring Customer Fluctuation using Secondary Data, Part One	Chapter 5: Measuring Customer Fluctuation using Secondary Data, Part Two	Chapter 6: Measuring Customer Fluctuation using a Cross-Sectional Survey	Chapter 7: Comparison of Measurement Approaches
I. To develop, measure and apply a new concept of market change analysis based on change segments	**RO1.** To understand conventional measures of market change in the fields of marketing and public transport and explore the benefits and drawbacks of these methods	✓	✓	×	×	×	×
	RO2. To develop a new concept for market change analysis based on change segment	✓	✓	×	×	×	✓
	RO3. To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne	×	×	✓	✓	✓	✓
	RO4. To explore the behavioural factors influencing market change segments using survey data for metro Melbourne	×	×	×	×	✓	×
II. To explore the potential of this approach as a means of improving market change analysis	**RO5:** To assess the potential of the new approach for application in industry	×	×	×	×	×	✓

3.3 Research Context

The development of a new concept to measure market change presented in this thesis uses the city of Melbourne as a case study. More specifically, research is focused on the area of Melbourne that is serviced by public transport. As such, it is useful that the research context is clearly understood, since it might influence the development of the customer fluctuation concept. It is anticipated that the customer fluctuation concept proposed might be applied to additional cities in the future, with minor adjustments to address changes in context.

The case for public transport in Melbourne is an important one that reflects the challenges of population growth, public transport and road infrastructure capacity and environmental concerns. These issues are also faced by many cities across the globe. Melbourne is the second largest city in Australia, just behind Sydney, with a population of approximately 4.9 million (as at 2017). It is also the Australian capital city experiencing the largest rate of growth, gaining 125,400 new residents in a single year (Australian Bureau of Statistics, 2018c). Growth projections conducted by the Australian Bureau of Statistics indicate that this growth is not slowing down, by 2021 the population is anticipated to exceed 5 million residents. Just 40 years after the 2011 census, the population of greater Melbourne is anticipated to be doubled, sitting at over 8 million residents (State of Victoria, 2016). A map showing Melbourne's and its public transport networks is provided in Figure 3.2.

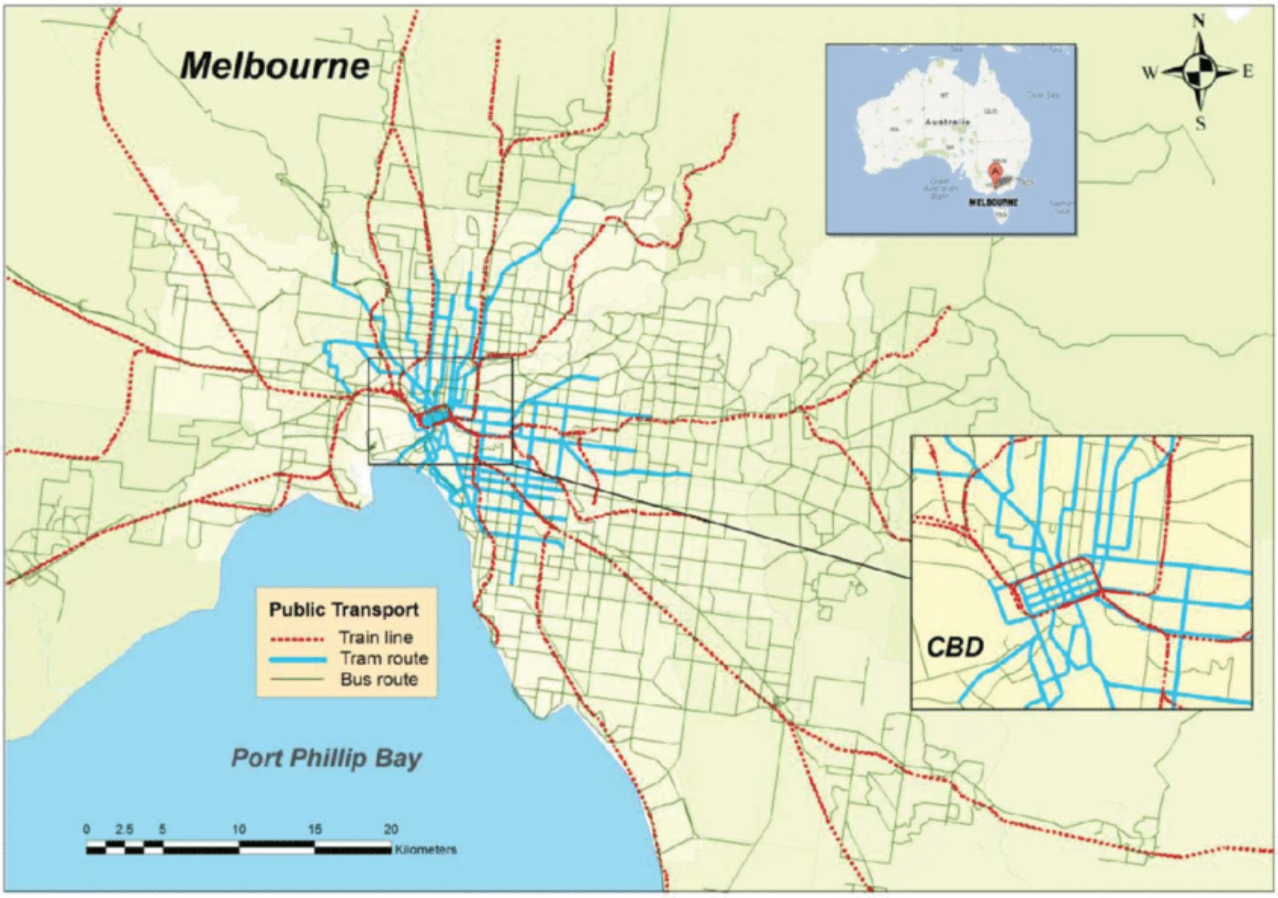

Figure 3.2 - Map of Melbourne Showing the Extent and Key Public Transport Networks (Duy et al., 2018)

Melbourne stretches outwards from a central business district located on the northern banks of the Yarra River, approximately 5 kilometres from the original Port, Port Melbourne. The river mouth is located on Port Phillip Bay. Typical of an Australian City, Melbourne is characterised by a sprawling city form and covers over 9992.5 km^2. From the CBD, Melbourne spreads by approximately 20km to the north, 30km to the east before being limited by the Dandenong Ranges, over 40km to the south and has ongoing sprawl across the western plains (City of Melbourne). The city centre is relatively dense and includes a mix of residential, commercial and retail uses. The inner city, within 10km from the CBD, is characterised by a medium density form with a large range of housing from low density heritage worker's cottages to mid- and high-rise apartment buildings. Beyond this, the built form of Melbourne is typically characterised by low density detached housing, including several greenfield developments in the outer most ring. Due to a sprawling urban form, the car remains the primary mode of transport for over half of Melbournians, as captured via the journey to work estimates (Cooper and Corcoran, 2018).

In Melbourne, public transport is overseen by the government body Transport for Victoria (TfV), also previously known as Public Transport Victoria (PTV). Melbourne's public transport system is provided across three key modes; the train, bus and tram (or more technically, streetcar) (Vuchic, 2007). The train is a large radial system that moves between the outer suburbs and the city and is currently operated by 'Metro'. Regional trains that travel from Melbourne to other Victorian cities are currently delivered by 'V/Line'. The inner city is supported by Melbourne's tram network, which takes a similar radial form between train lines. This tram network is currently operated by 'Yarra Trams'. The bus network services a much wider area of Melbourne as well as providing many east-west linkages that are missed by the bus and train network. The bus network is currently operated by 13 different operators, including the larger 'Trans Dev' and several smaller family-run bus companies. These smaller operators are represented by the Bus Association of Victoria. In total, there are 16 metro train lines, 24 tram routes (including the city circle) and 340 bus routes.

There are several types of tickets that are available for use within Melbourne's public transport system, called 'myki'. Melbourne has a smart card based ticketing system that generally requires users to 'touch on' and 'touch off' to record travel. myki cards suit a variety of user needs including full fare, concession fare and child myki tickets (Public Transport Victoria, 2018).

Melbourne's Public Transport has an integrated system with fares that are based on movement within or between two distinct zones (Zone 1 and Zone 2). These fares are calculated via an integrated system and do not vary by mode of transit used. This presents several interesting challenges when reviewing myki Data compared to traditional systems where separate tickets are required for each mode. An integrated fare system allows for movement between modes and routes within the system. As an example, all train trips require a touch on when entering the station and a touch off when exiting a station, but if switching trains without leaving a station, a user is not required to touch on and off again unless passing through ticketing barriers (Public Transport Victoria, 2018). Conversely, moving from the bus to the train would require two separate 'touch ons'. This can create limitations for the collection of data using the myki system as certain policies around fare collection mean that a touch on is not always necessitated and myki is not able to capture all trips taken by an individual user.

Each mode operates within its own unique set of challenges. In terms of patronage, Melbourne Train and Tram services generally continue to experience aggregate growth, while bus patronage is in decline. For the 2016 - 2017 reporting year TFV reported 236.8 million train passengers with 0.5% growth, 204.0 million tram passengers an increase of 0.2% and 118.0 million bus passengers a decrease of 4.0 per cent for the year (Public Transport Victoria, 2017). This decline in bus usage is a 5-year trend, although there was growth recorded on some bus routes. For all modes, the highest rates of ridership occur during peak commute times; 7 – 9am morning peak and 4 - 6pm.

In addition to ridership, authorities (such as TfV) also report on customer satisfaction, service reliability, service punctuality and fare evasion rates. This can provide some further information in regards to the variation in patronage rates between modes. A core example of the differences between modes is the rate of touch-ons, which represent roughly 90% of train boardings, 70% of bus boardings and 40% of tram boardings. This is influenced by different boarding approaches and levels of security between each mode. This context has been an important consideration for informing the development of a new concept for measuring market change – customer fluctuation.

3.4 Development of the 'Customer Fluctuation' Concept

The literature review presented in Chapter 2 identified a gap in public transport knowledge related to the analysis of market change. This research identified that there were four elements of market change that are routinely measured: acquisitions, defections, retention and internal variability. However, across both marketing and public transport disciplines, there was only a single instance of all four elements of market change being measured with a single tool (Mason et al., 2011). Most work focused on the interplay between two elements of market change, most commonly the relationship between defections (or churn) and retention, which fails to provide a full picture of market change.

The research of Mason et al (2011), did broadly consider all four elements of market change, however the predominant focus was acquisitions (new users), retention (retained users) and defections (lost users). Internal variability, or changing travel behaviour (such as starting, stopping and starting to travel again) was discussed though not implicitly measured by the framework used in the study. This identifies a clear gap in the existing literature for a single tool that actively measures all four elements of market change.

The growing topics of loyalty, market change and variability in individual public transport rider behaviour, show that this is an area of increasing interest. This research theorises that combining the four elements of market change into a single measurement tool is a possible way to improve our current understanding of, and approach to, market change analysis. **As such, a new approach, which measures customer acquisitions, retention and defections, as well as internal variability, using customer segmentation as the measurement approach will enrich the field of market change analysis.**

We term this approach, 'Customer fluctuation' which seeks to create disaggregate market segments. The concept captures the internal variability, or fluctuation, of ridership within public transport markets. For this Thesis, the concept of customer fluctuation can be defined as follows;

> *"**Customer Fluctuation:** A measurement concept that seeks to measure market change in time by separating markets into disaggregate segments based on changes in ridership (for example starting, stopping or continuing to travel) This concept measures the interplay between new, lost, retained and returning customers within public transport markets."*

This concept has been developed building on the marketing principles of customer churn and adaptions required to make it relevant to public transport markets. To better understand customer fluctuation, a description of the four segments based on ridership style are provided as follows:

- ***New riders:*** *This is to capture riders that are new to the transport mode, within the study period. Where possible, attempts will be made to measure first-time riders, but this is considered a small portion of the population (e.g. tourists).*
- ***Lost riders:*** *Those that have been consistently using a public transport mode and then stop using the mode and do not use it again within the period being studied*
- ***Retained riders****: Users that travel consistently each month with no more than 25% of the total study period as a break over non-consecutive months.*
- ***Returning riders:*** *Any user that rides sporadically in the period, consistently starting and stopping travel*.

The segments of new, lost and retained riders were based on the existing concept of customer churn as discussed by Riebe et al. (2014). In addition, the 'returning' category has been added to reflect the element of time and changing customer needs over time. This is based on the criticisms of churn posited by Tamaddoni et al. (2014) and Kamakura et al. (2005).

These segments are all a function of user decisions, to start, stop or continue using a public transport mode. A model illustrating the four segments is illustrated in Figure 3.3 and has been developed using the Dick et al.'s (1994a) four quadrant loyalty model, which looks at customer churn from the individual user perspective. For example, new users are those that decide to start and then continue to use public transport, whereas lost users are those that are continuing to use public transport and decide to stop.

		Decision Point	
		Start	**Continue**
Decision Point	**Continue**	New	Retained
	Stop	Returning	Lost

Figure 3.3 - A Four Segment Model of Customer Fluctuation Source: Authors concept based on Dick et al. (1994a)

3.4.1 Temporal Variables

A key component of measuring customer fluctuation is the question of the appropriate time period over which the assessment of segments is made. Customer fluctuation is a measurement that provides a summary of ridership behaviour over the identified measurement period. A key question in the development of the concept is the identification of an appropriate time frame to capture changes in individual ridership. The measurement period must be long enough to capture actual changes in ridership, rather than just seasonal variations, but not so long as to exaggerate change. For example, if smart card data was being used, a measurement period of four years would exaggerate the rate of lost users because smart cards tend to need renewal after a four-year period. All new card holders might be considered new users even though they had replaced a card so were more likely to continue their existing travel patterns. For initial explorations, a one-year measurement period was selected to reflect the existing standard reporting periods for public transport operators.

The other important temporal variable is the unit of time used to measure change across the year. Once again it is required to select a temporal unit with an appropriate level of sensitivity to the expected changes in behaviour over the measurement period. The issue of sensitivity is illustrated in Figure 3.4, through a conceptual analysis which shows that as the analysis period decreases the proportion of returning users will increase until all users would be identified as returning (stop/start travel).

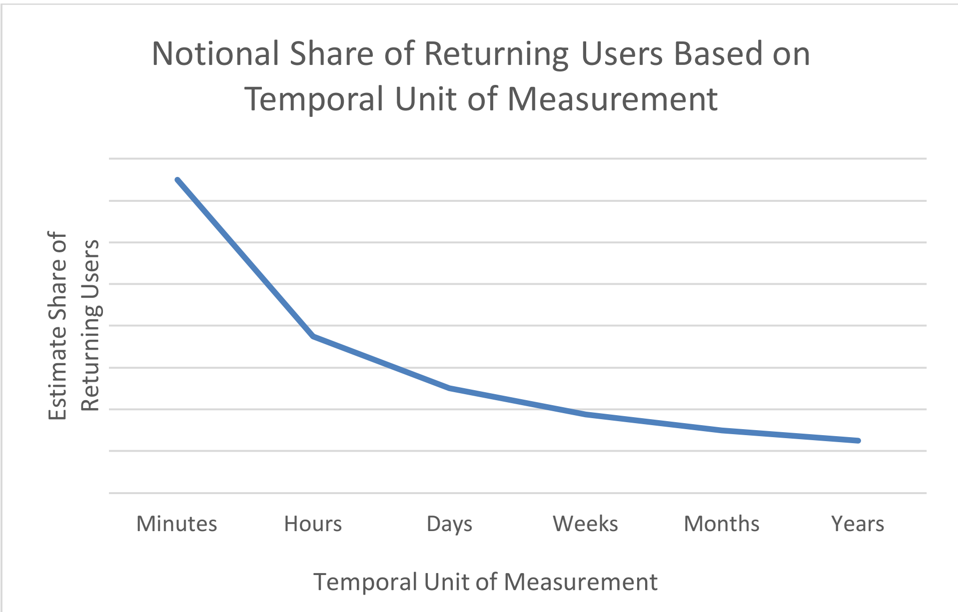

Figure 3.4 - Notional Plot Showing Impact of the Temporal Unit Selected on the Share of Returning Users

Based on this analysis, it has been identified that both months or weeks may be suitable as the temporal unit of measurement. It is suggested that both be explored in the approaches used to measure customer fluctuation. This is discussed further in Section 3.7 of this chapter.

3.5 The Rationale for Measuring Customer Fluctuation

Before identifying the suggested research approach for measuring customer fluctuation, it is important to review the rationale for its measurement. If customer fluctuation can be measured, it may allow for the setting of benchmark rates of acquisition, defection and retention within the public transport industry. This information could be used to identify if the degree of observed customer defections and acquisitions is unusual (Riebe et al., 2014). This is useful information for service providers, as an unusually high rate of defections (above the benchmark) may be an indication of customer dissatisfaction with the service. It can also allow for comparisons to be made with competitors. Measuring customer fluctuation relies on a similar approach to market churn, predominantly focused on measuring the customers leaving a market and comparing it to the amount of those retained, new or returning within a designated period.

The measurement of customer acquisitions, defections and retention have been conducted for a broad range of service markets (Bagchi and White, 2005), although seldom are all three reviewed holistically due to competing theories of market growth (Riebe et al., 2014). Subscription based service markets that rely on repeat purchasing behaviour (for example, phone plans or credit cards) have been the predominant focus for such studies (Ahn et al., 2006). Public transport markets similarly rely on repeat purchasing behaviour, though without a formal obligation for repeat use. Yet, although the industry currently measure patterns of travel, we do not know the exact number of customers that have commenced, ceased or continued using public transport over time in the case of Melbourne. If operators seek to understand patterns of public transport ridership to appropriately cater for, grow and predict demand for services, a tool that can measure and review the interplay between customer acquisition, retention and defections would be of value. Developing a practical tool that can account for all variations in market behaviour holistically can also assist in our ability to understand and compare related markets of different sizes.

The measurement of public transport ridership markets is generally limited to aggregate market data, identifying only whether there is net growth, decline or stability. This limits the ability to understand and grow markets. It is likely that this has occurred due to historical limitations in the availability and cost of collecting data. The addition of new data sources, such as online survey collection and smart card data has opened up new options for exploring travel data. Morency et al (2007) illustrates the potential of smart card data to measure the day-to-day variability in transit use but identifies a need for more user friendly tools to do so. Also identified are limitations in existing approaches to measure travel patterns, namely the difficulty in capturing 'non-typical patterns', particularly caused by a lack of detail for each individual rider (Yoh et al., 2012, Morency et al., 2007). The development of a framework that will measure the number of customers that commence, continue, cease and re-commence travelling with a public transport mode over time will contribute significantly to our existing understanding of transport markets. Not only that, but such a tool can provide us with an understanding of the influences not just behind decisions to cease or not use a service, as is the current academic focus, but also to commence or continue use. This has been recognized as a valuable contribution that may improve the accuracy of existing models (Ma et al., 2013, Chu, 2015).

To further illustrate the value of measuring customer fluctuation within public transport markets, Figure 3.5 illustrates the way that new, lost, retained and returning users interact within a market.

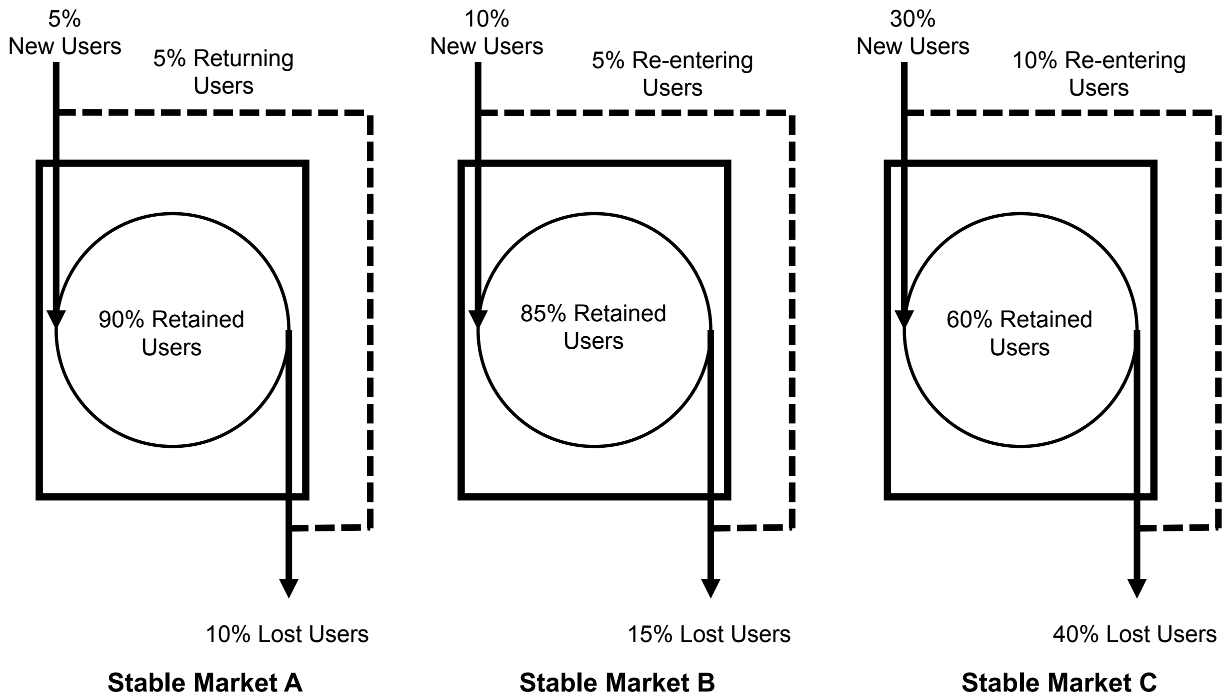

Figure 3.5 - Customer fluctuation within a Stable Market Context adapted by the author from Blythe (2009)

Based on current methods of measurement, all three of these markets pictured in Figure 3.5 would be considered the same i.e. stable markets exhibiting no growth or decline. However, each of these markets are different, becoming increasingly more porous. As an example, Mason et al. (2011) measured customer churn using a broad survey of users, with a more detailed panel survey to detect the influence of churn over a period of two years. However, this work divided users into commuter and leisure travel and modified measurements accordingly and does not provide an understanding of the entire market.

Simply understanding net market change does not provide a detailed understanding of market change and the behaviour of people within the market. Measuring ‘customer fluctuation’, or significant changes in the frequency of use (primarily commencing or ceasing use) of individual users in a market, in this case public transport markets, will strengthen our understanding and ability to grow markets. However, to develop an appropriate research approach, it is essential to first investigate the existing tools for measuring public transport markets within Melbourne that may be able to measure customer fluctuation.

3.6 Existing Measurement Tools

For Melbourne, the focus of this Thesis, travel is currently measured using several different sources, including;

- An annual report that reports on annual patronage, customer satisfaction, service reliability, service punctuality and fare evasion rates;
- A Household Travel Survey (VISTA); and
- An authority lead ‘tracker’ survey.

These sources were reviewed to assess their potential to be adapted to measure customer fluctuation. The annual report was provided data which was very aggregate in nature. It did not provide information around individual ridership or changes in individual rider habits. The Household Travel Survey (VISTA), although a valuable resource, provides information for only one survey day of travel over numerous years and is not limited to the same day for all respondents. Although the survey identifies whether users are travelling more or less than the previous year, it does not provide sufficient longitudinal data or clarity to aid in the measurement of customer churn.

The tracker survey was identified as the existing measurement tool with the highest potential for capturing customer fluctuation segments and individual changes in the rate of travel. A summary of the assessment of this approach for its ability to capture customer fluctuation is provided below.

3.6.1 Tracker Survey

The tracker survey samples 400 transport users within the Melbourne Metro area each quarter (133 respondents each month) using a random telephone survey. The survey provides valuable data as it collects responses across all public transport modes allowing for comparison. Data is collected monthly and analysis was conducted on the data collected between October 2013 and November 2015.

Data is collected through random sampling of Melbourne households. This allows for the capture of infrequent users more effectively than alternative methods, such as on-board or intercept surveys. The questionnaire is focused on public transport use with the flexibility to add questions dependant on current authority research interests. Based on the review, the tracker survey presents several benefits for the measurement of customer fluctuation:

- It is conducted monthly within the Melbourne Metro area from a panel of public transport users.
- It is of a sufficient sample size to provide a 95% confidence interval +/- 5% for the Melbourne metropolitan area.
- It collects data on individual user travel habits for each public transport mode.
- It collects a sample over multiple years, which could allow for the identification of trends and new users.
- It identifies a comprehensive list of reasons why people make the travel decisions they make, including reasons for why people have chosen to enter or exit the market.

A review of tracker data for this thesis has determined that without the addition of a new block of questions, or significant alteration to the wording of existing questions, the tracker survey would not be able to measure customer fluctuation. This was because:

- The tracker survey does not ask individuals to identify the number of times they travel on each mode for any period, meaning it is impossible to identify new and lost users.
- The wording of questions relating to user frequency relied on imprecise wording (‘a little more’, ‘about the same’, ‘a little less’) that could not be accurately divided into fluctuation categories.
- The use of a year as the temporal unit of measurement was not considered sufficiently fine-grained to identify changes in frequency and the nuances of customer fluctuation over time.
- The sampling for this survey varies each quarter and does not track the same people over time.
- The survey does not include people that have previously used public transport but have since stopped.

Overall, the data collected through the tracker survey, though valuable, cannot directly measure customer fluctuation, individual travel patterns or small-scale (week or month) changes in travel patterns. **Given these limitations, the tracker survey was identified as an inappropriate tool for measuring customer fluctuation.** Instead, the findings of this assessment were applied to the development of a new framework, including an online survey that could more readily capture customer fluctuation.

3.7 Research Approach

As there are no existing tools that could measure customer fluctuation, a mixed method approach has been adopted to explore the application of this new concept to public transport markets. The research approach has focused on quantitative information to investigate the measurement of customer fluctuation. The data used includes a mix of secondary data, collected for other purposes, and primary data collected for the exclusive use of this study. This mixed method approach is used to ensure the research findings are robust and account for the limitations associated with a single approach.

An overview of the research approaches used is provided in Table 3.2.

Table 3.2 - Overview of Research Approach

	Chapter 4: Measuring Customer Fluctuation using Secondary Data, Part One	Chapter 5: Measuring Customer Fluctuation using Secondary Data, Part Two	Chapter 6: Measuring Customer Fluctuation using a Cross-Sectional Survey	Chapter 7: Comparison of Measurement Approaches
Type of Analysis	Quantitative Exploratory Analysis	Quantitative	Quantitative	Quantitative
Research Objectives Addressed	**RO3:** To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne	**RO3:** To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne	**RO3:** To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne **RO4:** To explore behavioural factors influencing market change segments using survey data for metro Melbourne	**RO2:** To develop a new concept for market change analysis based on change segment **RO3:** To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne **RO5:** To assess the potential of the new approach for application in the industry
Method	a-priori segmentation of smart card data and exploration of data	a-priori segmentation of smart card data and exploration of data	Cross-sectional panel survey	Analysis of all results

It is worthy of note that research objective one (RO1) was addressed through the exploratory literature review that has been presented in Chapter 2. A summary of the methods used now follows, more detailed information is provided in the applicable chapters where necessary.

3.7.1 Secondary Data Assessment, Part One

The key aim of the secondary data assessment is to explore the identified approach for the measurement of customer fluctuation segments through the application of the approach to smart card data. Smart card data was identified as a key resource for measuring the travel patterns of individual riders. This was based on the growing use of smart card data in the surrounding literature and practice (e.g. Chu, 2015, Ma et al., 2013, Zhang et al., 2018).

Smart card data for this thesis was obtained from Transport for Victoria (TfV). This included a random sample of cards travelling between October 2016 and September 2017. The data available provided 40,850 smart card IDs for the bus, 47,580 for the tram and 84,379 for the train. Melbourne smart cards are used for all modes. Some smart cards may be counted across all three modes, while some may appear for only one mode. The data provided included each individual entry (comprised of a touch-on and the corresponding touch-off) for the sample of smart cards over the entire year. To simplify the data analysis, this complex system was reorganised using a purpose-built coding process in Python. This provided a similar function to the density-based spatial clustering of application with noise (DBSCAN) and K Means + algorithms used in previous studies (Ma et al., 2013). However, due to the intent to measure a new concept, a purpose-built code was determined as the most efficient and practical method for reformatting and measuring the data provided.

It is assumed that each individual smart card represents a single user and that use of the card during the period of analysis represents their travel behaviour. The code used was designed to isolate each individual smart card ID, identify each appearance over the year, and summarise the number of individual trips completed by that ID for each month or week over the measurement period. The code also removed extraneous data that was not being used in the smart card data analysis (for example, the stop ID, station name, station coordinates).

Figure 3.6 illustrates the changes in the data format that were obtained through this process, going from the data set as provided, to the data set used for the a-priori segmentation of users into customer fluctuation segments.

Card ID	Trip Date	Mode	Station/ stop ID	Station/Stop Name	Co-ordinates	End Trip Date and Time	End Trip Station/ Stop ID	End Trip Station/ Stop Name
12345678	2017-06-05	Metro Bus	9707	Trickey Ace/Overton Lea BVS (Sydenham)	-37.696806 144.765989	2017-06-02 16:09:36.00	9712	Hume Dr/Overton Lea Bvd (Sydenham)
12345679	2017-10-19	Metro Train	19945	Armadale Railway Station (Armadale)	-37.856452 145.019328	2017-10-19 11:37:52.000	64404	Flinders Street Railway Station (Melbourne City)

Data restructure through Python Coding

Card ID	Oct '16	Nov '16	Dec '16	Jan '17	Feb '17	Mar '17	Apr '17	May '17	Jun '17	Jul '17	Aug '17	Sep '17
Bus Users												
12345678	15	11	10	0	0	0	0	4	3	0	0	1
Train Users												
12345679	0	0	0	0	0	2	5	7	0	0	0	0

Figure 3.6 - Process of reformatting data for Measuring Customer Fluctuation

Following the re-organisation of data, the next step is an a-priori segmentation process.

3.7.1.1 A-Priori Approach for Passenger Segmentation based on Fluctuation Principles

The initial approach to the market segmentation analysis adopts an a-priori segmentation process where passengers are classified in accordance with their temporal patterns of travel. In the a-priori market segmentation, clusters are defined by pre-selected characteristics chosen by the researcher (Elmore-Yalch, 1998, Kieu et al., 2015b, Kieu et al., 2014). The results of the segmentation do not influence the characteristics of the predefined segments. An a-priori segmentation assumes that although individual ridership is highly variable, there are four distinct classes of transit riders, as outlined by the new concept of 'customer fluctuation' – 'new', 'lost', 'retained' and 'returning'. This approach has also been termed 'physiological segmentation', as it is the type of travel patterns exhibited that determine the segment (Elmore-Yalch, 1998).

As customer fluctuation is a new concept, several issues had to be considered, including the appropriate measurement period and temporal characteristics of each group. This initial application seeks to explore different definitions and approaches found in the literature to identify a simple and effective tool for measuring changes in an individual's public transport ridership. The initial test adopts one year of smart card data (between October 2016 and September 2017). As identified in the development of the concept, both months and weeks are tested as the temporal unit of measurement.

Table 3.3 A-Priori Segmentation Rules for Measuring Customer Fluctuation with Coding Parameters

Segment	Description	Coding Parameters
New Users	Where the first active trip occurs after December '16 and there are continuous trips each month with no more than one month/four weeks break the user classifie**s as a 'new user**.'	After the first 2 months
	Where a smart card/user appears in the last two months of the study period (September or October), they **classify as 'new users'**. This rule includes users that appeared in September but not October, as they did not yet qualify as 'lost users' due to a lack of time for consecutive misses.	Last 2 months of the measurement period
Lost User	Those that have several consecutive travel months and then miss two or more consecutive months and do not return in the measurement period classif**y as a 'lost user.'**	Miss more than 2 consecutive months and do not return
Retained user	Where users have a minimum of one active trip per month for the entire period, with no more than two consecutive months break the user classifie**s as a 'retained user**'. These breaks are to allow for travel/annual leave	One active trip per month, no more than 2 consecutive months' break
	Smart cards/users were still identified as **'retained user**s' if they had multiple breaks in travel if these breaks did not occur in consecutive months and accounted for less than 25% of the overall period.	Missed less than 3 months' overall
Returning User	Smart cards/users that had returned once were assumed likely to return and thus were not categorised as lost if they had two consecutive months of no travel after travelling and returning to travel within the measurement period. These users remained classified as a **'returning user'.**	n/a
	If there is greater than 25% of the measurement period without travel (consecutive or non-consecutive) followed by travel, the smart card/user will be a **"returning user"**.	Missed more than 3 months' overall
	Those that have an irregular travel pattern inconsistent with the above definitions are known as a **'returning user'**.	n/a

Each smart card was assumed to represent a single public transport user. It was then segmented into customer fluctuation categories, based on the a-priori segmentation rules provided in Table 3.3.

Several assumptions have been utilised for the development of the a-priori segmentation rules described above. These assumptions are briefly summarised as follows:

- An active trip is defined as a single journey within the temporal unit of measurement, for a user to be identified as inactive they must take 0 trips within the temporal unit of measurement.
- All users in the data had a minimum of one trip during the measurement period.
- The measurement period starts in mid-October '16 and ends in mid-October '17. Months are counted from this point and do not follow a standard calendar month.
- New users are required to maintain a retained ridership pattern following their first appearance. This is considered appropriate due to the reduced time they are present and the allowance for retained users to miss no more than 2 consecutive months, or 25% of the measurement period over all.

- The two months without travel is adopted to classify users as lost or returning. This effectively provides a window of almost 3 months for a user to travel a minimum of once before they are confirmed as lost or returning. By travelling on the last day of the third month, they would remain a new or retained user. The leniency in this definition is to allow for the appropriate classification for users that travel at different frequencies.
- Finally, all users that did not meet the parameters for new, lost or retained users were classified as returning

The assumptions identified for each behaviour change segment (new, lost, retained, returning) were devised to allow for users to undertake standard variations in travel over the course of the year, without necessarily influencing their ridership segment. For example, retained users could not travel for 25% of the overall period while remaining retained to allow for the impacts of school holidays, annual leave, illness or other regularly occurring factors likely to occur within the measurement period. A visual representation of the above rules can aid interpretation (Figure 3.7 and Figure 3.8). This depiction provides a review of typical and non-typical travel examples, but does not account for all possible variations.

Month	Nov ‘16	Dec ‘16	Jan ‘17	Feb ’17	Mar ‘17	Apr ‘17	May ‘17	Jun ‘17	Jul ‘17	Aug ‘17	Sep ‘17	Oct ‘17
New												
Lost												
Retained												
Returned												

Key: Months where user travelled at least once

Figure 3.7 – Typical Examples of Customer Fluctuation Segment Patterns based on A-priori Rules

Month	Nov ‘16	Dec ‘16	Jan ‘17	Feb ’17	Mar ‘17	Apr ‘17	May ‘17	Jun ‘17	Jul ‘17	Aug ‘17	Sep ‘17	Oct ‘17
New												
Lost												
Retained												
Returned												

Key: Months where user travelled at least once

Figure 3.8 - Less Typical Examples of Customer Fluctuation Segment Patterns based on A-priori Rules

The Customer Fluctuation segmentation process for smart card data is summarised in Figure 3.9.

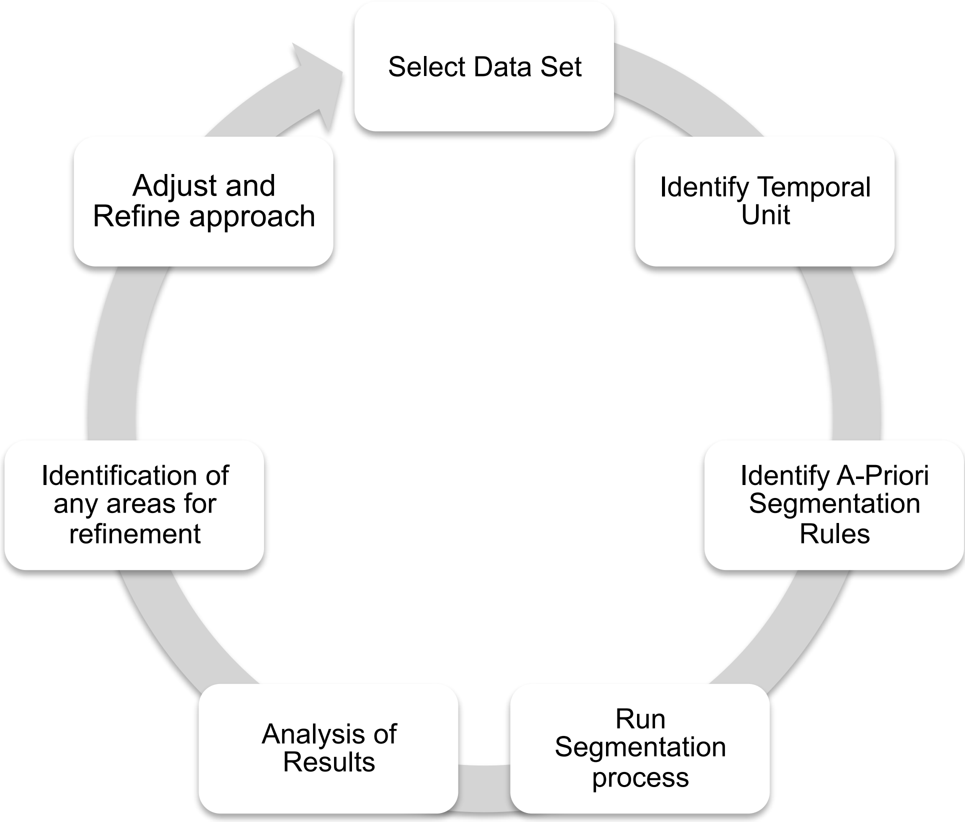

Figure 3.9 - Exploratory Customer Fluctuation Segmentation Process with Smart Card Data Adopted in this Research

Figure 3.9 highlights the iterative process required to explore the concept of customer fluctuation. This includes the identification of the relevant data set and the appropriate temporal unit. Following this the a-priori segmentation rules are identified based on similar studies in the literature and observations around patterns and trends in travel behaviour. Once the first segmentation process has been completed, the results are analysed to find areas for further adjustment and the approach is refined and repeated. All changes will be clearly noted in the relevant chapters (4 and 5).

3.7.2 Primary Data Assessment

As this is an exploratory study, the use of cross-sectional primary surveys was adopted, building on surveys adopted by Mason et al (2011). The aim of this survey was to explore a different approach to measuring market change segments. In addition, as a primary survey it was possible to ask users why they were in different segments.

The customer fluctuation survey was developed as an online survey that samples both public transport users and former public transport users within the Melbourne Metropolitan Area. The survey targets users from all modes (train, bus and tram). Although traditional cross-sectional survey's fail to capture the temporal element of measuring behaviour change (Saleh and Farrell, 2007, Behrens and Mistro, 2010), a retrospective approach has been utilised to capture longitudinal information through a single survey (Beige and Axhausen, 2008). Questions were simplified to reduce any potential issues with memory recall. This included asking participants to identify their 'mainly used mode' and to recall whether they had travelled a minimum of once for each month of the previous year. Participants were then asked to categorise their travel based on a series of descriptive statements. Finally, as a core focus of the survey is to identify the factors/determinants behind customer fluctuation behaviour, they were asked to identify the top three reasons that influenced their choices.

The survey was managed and supplemented with data collected by a commercial market research company (IPSOS) and the research team. The survey platform 'Qualtrics' was used to facilitate and store responses. Data collection occurred between November 2018 – February 2019. Participants that responded through the market research survey link, were provided with a small payment for their time in completing the survey. This was completed independently from the research team. A sample frame was developed to target a sample that is broadly representative of the population of Melbourne. It was structured by age, education and income, using response quotas.

The survey was kept short to encourage a higher response rate and limit the imposition on respondents. The survey was piloted several times with a small sample of individuals, including academic and non-academic participants, to check for clarity and efficiency of the information being collected.

3.7.2.1 The Survey Sample

The survey sampled Melbourne residents, with a focus on those that currently or recently used public transport. However, a small number of non-public transport users were also targeted to attempt to reach a proportion of lost users, as an attempt to address response bias. The sample aimed to estimate the population of customer fluctuation within a degree of statistical accuracy. The sample size for the survey was calculated using the equation in Figure 3.10.

$$\text{Sample size} = \frac{\frac{z^2 p(1-p)}{e^2}}{1+\left(\frac{z^2 p(1-p)}{e^2 N}\right)}$$

Where N = population size, e = Margin of error (percentage in decimal form) and z = z-score

Figure 3.10 - Power Equation for Determining Appropriate Sample Size Source: (Taherdoost, 2017, Survey Monkey, 2019)

This calculation was used to determine the sample size required for both 99% and 95% accuracy both +/- 5. This used the ABS record of Melbourne's population as at 30 June 2017, which was

4,850,740 (Australian Bureau of Statistics, 2018b). A sample for 99% accuracy (+/- 5%) is 666 respondents and for 95% accuracy (+/- 5%), 385 respondents would be required. As this survey sought to compare three different modes and would need each mode surveyed to be statistically significant, it was identified that this would need to be repeated for all three modes. This identified a target sample of 1,998 respondents for 99% accuracy and 1,155 respondents for 95% accuracy for each mode. In the instance of low response rates, a sample size for 90% accuracy (+/- 5% was also calculated), this would require a minimum of 269 participants per mode. This is summarised in Table 3.4.

Table 3.4 - Calculations of Required Sample Size for Melbourne Metropolitan Area

		One Mode	Three Modes
Statistical Significance of Sample	90% (+/- 10%)	269	807
	95% (+/- 5%)	385	1,155
	99% (+/- 1%)	666	1,998

Due to the financial and time limitations of this project, a 95% accuracy +/- 5% was the statistical significance target adopted for survey responses.

3.7.2.2 Survey Structure and Questionnaire Development

This section details the survey structure and questionnaire development. The structure of the survey is illustrated in Figure 3.11; the complete questionnaire is provided as an appendix to this thesis.

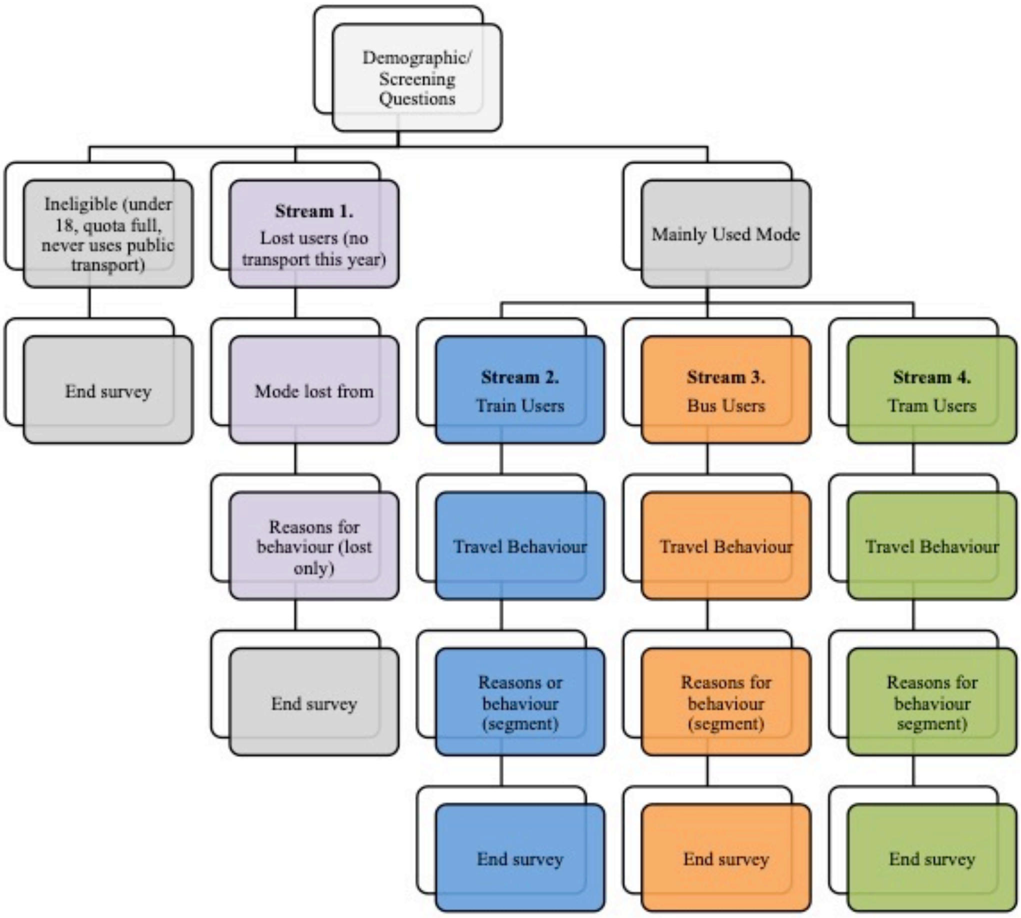

Figure 3.11 - Diagram of Questionnaire Structure Highlighting the Four Streams

Prior to commencing the survey, participants were first taken to an explanatory statement and asked to provide their consent to participate. Once they had agreed to participate they were asked a series of demographic and screening questions. Screening questions were included to ensure responses reflected the desired sampling frame. The sampling frame used sought to align with the demographic characteristics of Melbourne based on the ABS census data (Australian Bureau of Statistics, 2018a), while capturing a minimum of 80% current public transport users and a minimum of 5% non-public transport users. This requirement was to ensure the survey would be able to capture lost users, while also reducing the risk of response bias from invested public transport users. The sampling frame used is provided in Table 3.5.

Table 3.5 - Desired Sampling Frame for Customer Fluctuation Survey

	Melbourne (%)	Participation Quota (target 1,200 participants)
Gender		
Male	49%	588
Female	51%	612
Other	-	
Age		
18 - 19	1.5	18
20 - 29	15.5	186
30 - 39	15.5	186
40 - 49	13.9	167
50 - 59	11.9	143
60 - 69	9.3	112
70+	9.7	116
Employment		
Employed full-time	58	696
Employed part-time	30.6	367
Unemployed looking for work	6.8	82
Student	-	
Retired	-	
Home duties	-	
Other	4.6	55
Household Structure		
Family with children under 18	50.3	604
Family with adult children	25.3	304
Couple with no children	23.1	277
Single	23.2	278
Group	5.0	60
Other	1.5	18
Income		
Negative/nil income	-	
$1 - $299 ($1 - $15,548 p.a)	3.4	41
$300-$499 ($15,600 - $25,948 p.a)	8.6	103
$500 - $799 ($26,000 - $41,548 p.a)	11.0	132
$800 - $1,249 ($41,600 - $64,948 p.a)	15.5	186
$1,250 - $1, 749 ($65,000 - $90,948 p.a)	14.8	178
$1,750 - $2,999 ($91,000 - $155,948 p.a)	26.4	317
$3000 or more ($156,000 or more p.a)	17.8	214
Prefer not to say	2.5	30
Education		
Post Graduate Degree Bachelor Degree	37	444
Graduate diploma or certificate	12.9	155
Year 10 or above	29.6	355
No educational attainment	1.2	14

*Melbourne Data from the 2016 Census of Population and Housing Community Profile for the Greater Melbourne Area (Australian Bureau of Statistics, 2018a)

The sampling frame (Table 3.5) was used to create participation quotas, using the target of 1,200 responses (400 per mode). Where quotas were exceeded, participants were not eligible to participate. However, quotas were relaxed to increase representation where users were part of an under-represented demographic group or mode. During sampling, requirements were adjusted to achieve the best sample possible within the time frame. As an example, the number of respondents that mainly used the bus was low and so quota limitations were removed for bus users.

Once demographic questions were answered, the remainder of the survey was divided into four separate streams; lost users (all modes), mainly train users, mainly bus users and mainly tram users. Participants were directed into a stream based on questions around their main mode used and, if they no longer used public transport, whether they previously used it. The use of the four streams was selected to reduce survey time or the need to require repeat answers from participants.

A series of screening questions were asked to determine the appropriate stream for a participant to complete. This was also used to assist with managing survey quotas and ensuring participants met the required criteria for participation. Users were excluded from completing the survey if they were under 18, outside of the study area or quotas were full. Once users had been assigned to one of the four survey streams, they were then asked a series of questions to identify their customer fluctuation behaviour.

3.7.2.3 Identifying Customer Fluctuation Behaviour

This section of the survey focused on identifying participants travel behaviour over the last year. To identify new users, the first question asked whether users had travelled by that public transport mode in the following year. This was followed by a question asking users to identify the frequency within which they use public transport within a regular month.

The key question in this section required participants to recall their last year of travel. This was completed by requiring participants to select yes or no when asked whether they had travelled by their selected mode at least once for each month of the measurement period. The frequency of travel was omitted from this question to assist with ease of recall and was captured in later questions.

Participants were then asked to review their responses for travel per month over the last year and identify the statement that best reflected the travel pattern. These statements and their associated segment were provided as follows:

1. *I started using the train/tram/bus a few months into the year and used it most months once I started* **(new user)**
2. *I used the train/tram/bus for at least one month this year and then didn't use the train/tram/bus again* **(lost user)**
3. *I used the train/tram/bus every month, or most months* **(retained user)**
4. *I used the train/tram/bus sometimes* **(returning user)**

Users were not told when categorising themselves which segment they would be assigned to, to reduce selection bias. Once users had selected the relevant category, they were asked the applicable reasons to explain their customer fluctuation behaviour.

3.7.2.4 Identifying Reasons for Customer Fluctuation Behaviour

The second focus of this survey is to establish why participants adopted types of customer fluctuation behaviour. To create a framework for participant responses around reasons, four categories were provided based on a review of the literature surrounding factors influencing individual travel behaviour. These categories were:

1. Life Event
2. Lifestyle Choices
3. The Train/Bus/Tram Service
4. The Trip Taken

Users were required to select three influential reasons across all categories. Table 3.6 summarises how each category has been defined and examples of the options available for participants to respond to each category. Again, these were determined from an exhaustive review of factors influencing travel synthesised from previous research.

Table 3.6 - Summary of Framework to Identify Reasons for Customer Fluctuation Behaviour

	Definition	Examples
Life Event	Any significant changes, likely once off, in the participant's life circumstances.	Moving home or work locations, starting a new relationship, having children, leaving school or university, returning or going on an extended holiday and retiring
Lifestyle Choices	This reflects ongoing lifestyle decisions made by an individual, and includes both active choices made by the user and passive choices that have been required due to user circumstance.	Active lifestyle includes choosing public transport to save money or as an active environmental choice Passive lifestyle choices are indirectly made such as not owning a car.
The Train/Bus/Tram Service	This includes direct service attributes that are directly within service providers control.	Examples of service attributes include efficiency, frequency, cost and cleanliness.
The Trip Taken	This category reflects where the public transport mode has been selected based purely on its ability to serve a specific trip.	Examples include travelling to an event

The relevant literature sourced for each of the four categories are now outlined.

<u>Life events</u> influencing travel habits have been both a focus and a finding of public transport literature (Beige and Axhausen, 2012). Life events are taken to mean any significant changes in the participant's life circumstances. Examples of significant life events include moving home or work locations, starting a new relationship, having children, leaving school or university, returning or going on an extended holiday and retiring. Life events are included to reflect the assumption that travel behaviour is habitual, and habits may be broken by significant changes in a person's life circumstances (Verplanken et al., 2008, Clark et al., 2014).

The next category was **lifestyle choices**. This category is intended to capture either active or passive lifestyle choices. Active lifestyle choices are focused on personal values and such as choosing public transport to save money or as an active environmental choice (Abrahamse et al., 2009, Bamberg et al., 2007). As identified by Bamberg et al. (2007), these lifestyle choices can also capture the influence of wider societal norms. Passive lifestyle choices are those that have been made indirectly due to user circumstances. This reflects the concept of captive public transport users, or those who don't have access to alternative modes such as a car (short or long term) that they would use as a preference (Jacques et al., 2013).

The train/bus/tram service was identified as a key category due to the established link between service satisfaction and loyalty within public transport literature. Examples include studies of satisfaction with service and loyalty (van Lierop and El-Geneidy, 2016). Service attributes are of interest to public transport operators as they fall within the service providers control and therefore are the most direct factors to address.

The final category is the **trip taken**. This category is related to the specifics of the trip/s and how that has influenced travel behaviour. Trip specific behaviour includes the reason for travel (for example travelling to an event), the time of travel (I feel unsafe using public transport in the evening) or other external circumstances impacting travel (parking was difficult, the weather wasn't suitable). This reflects studies that show even habitual/regular travel can be subject to change when confronted with new information, such as changes in timetabling (Bamberg et al., 2003). There is also some potential that the increase of real-time data available for public transport modes, might impact on travel decisions, for example, catching the bus which comes a few minutes earlier than the tram (Buehler et al., 2017).

A detailed list of all reasons given by category is provided in Table 3.7

Table 3.7 - Reasons for Customer Fluctuation Behaviour by Category

Category	Reasons for Starting	Reasons for Stopping	Reasons for Continuing
Life Event	I changed home or work locations	I went on holiday	I changed home or work locations
	My life circumstances changed (I returned to work after having kids, I started a new job)	I stopped studying/finished school	I started studying
	I returned from holiday/ sabbatical	My life circumstances changed (e.g. I had children, retired)	My circumstances haven't changed (e.g. I take my kids to school on the bus)
Personal Factors	I am trying to save money on my transport costs	I no longer needed to make trips with the train/bus/tram	I believe it is important to take sustainable transport modes
	I was too sick or injured to drive or cycle	I was sick or injured and couldn't catch the train/bus/tram	I don't have a car or can't drive
	I do not own a car or can't drive	Catching the train/bus/tram was too difficult with children	I am trying to save on my transport expenses
	My car or bike was unavailable	I bought a car	I was too sick or injured to drive
	I believe it is important to take sustainable transport modes	I don't like the train/bus/tram	I regularly use all modes of public transport
	I have started a new hobby/ socialising more		
The Service	I enjoy travelling by train/bus/tram	Catching the train/bus/tram takes too long	The train/bus/tram is the most convenient option for me
	I find catching the train/bus/tram reasonably priced	Catching the train/bus/tram was too unreliable	I enjoy catching the train/bus/tram
	I feel safest when catching the train/bus/tram	The train/bus/tram route I was using changed or stopped	I feel comfortable catching the train/bus/tram
	I find the train/bus/tram to be less crowded than other modes	I felt uncomfortable on the train/bus/tram I was using	I like that the train/bus/tram is not too crowded
	I find the train/bus/tram to be reliable	The train/bus/tram wasn't available at the times I wanted to travel	I find the train/bus/tram reliable
	The train/bus/tram timetable changed to suit me better	The train/bus/tram was too crowded	I think the train/bus/tram provides a good service
		Catching the train/bus/tram got too expensive	
The Trip Taken	I used the train/bus/tram when traveling to an event	I travel at night and didn't feel safe	I routinely make the same trip and like knowing what to do
	Parking is too difficult at my destination	The train/bus/tram wasn't needed as part of my journey anymore	Parking is difficult at my regular destination
	The train/bus/tram is the most convenient option for the main trips I was making	I found it easier to drive to my destination	I get the train/bus/tram as just one part of my regular journey
	I had an unpleasant experience with a different mode of transport		
	I catch the train/bus/tram when the weather isn't suitable for other travel		

The data analysis in Chapter 6 primarily relies on using the statistical package SPSS. Descriptive analyses provide further insights into the quantitative data.

Participants were also provided an opportunity to provide any further comments. **The questionnaire is provided as Attachment A to this thesis.**

3.7.3 Comparison of Measurement Approaches

The subsequent chapters present the results of the analysis using the approaches described above. Smart card analysis results are presented in Chapters 4 and 5, with the primary survey results appearing in Chapter 6. This mixed method approach has been utilised as a form of triangulation. Triangulation is grounded in the principle that any isolated method of obtaining data has weaknesses (Connidis, 1983) and that stronger inductive reasoning can be performed by converging different forms of information about the same phenomenon (Denzin, 1978, Jick, 1979).

3.8 Conclusion

This chapter has provided a review of the context for this study, the development of a customer fluctuation approach and an overview of the mixed methods research approach selected for exploring the application of customer fluctuation to public transport markets.

The following chapters (4 – 7) explore these research methods in further detail.

Part B: Concept Testing

Chapter 4: Measuring Customer Fluctuation using Secondary Data, Part One

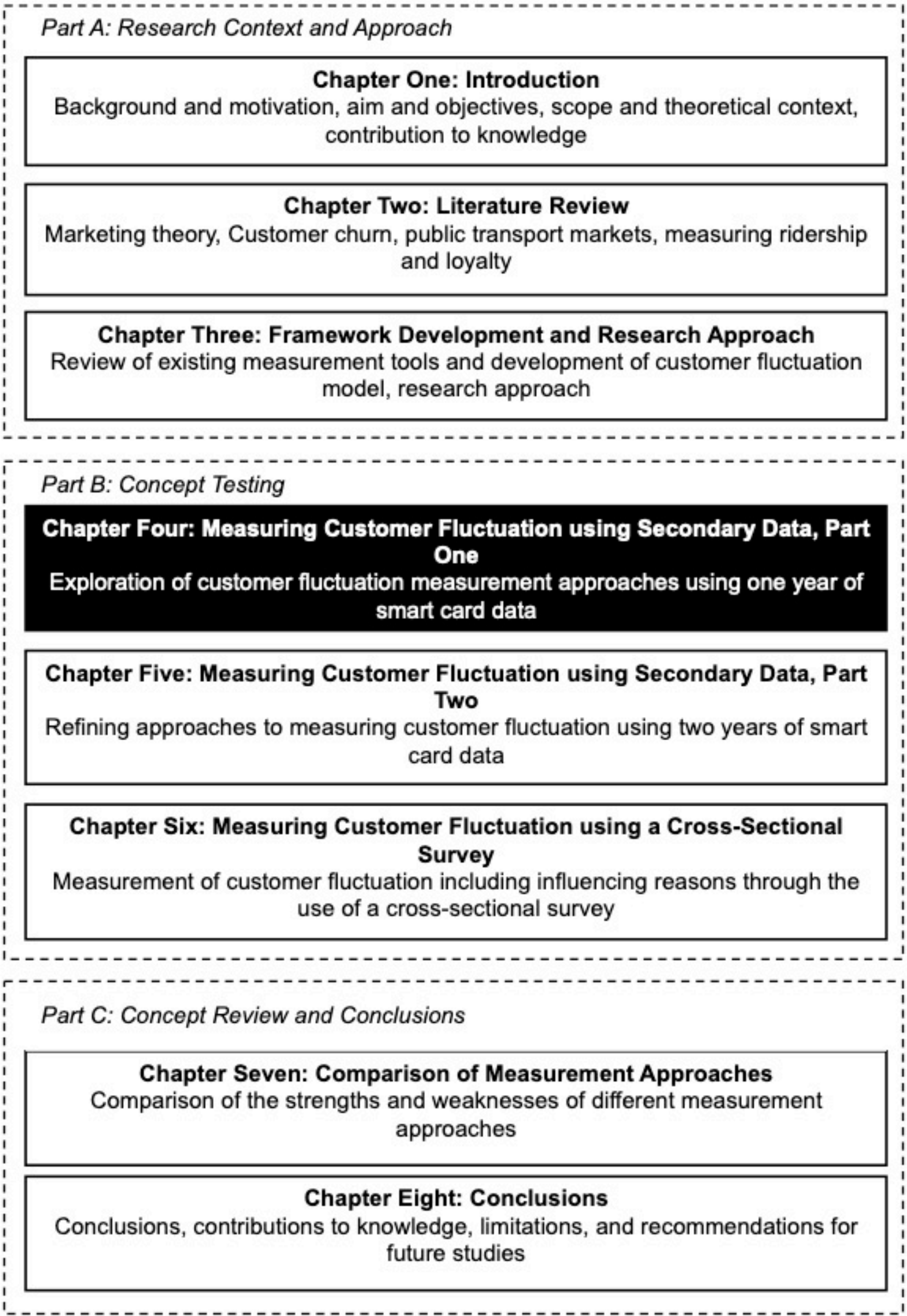

Figure 4.1 - Position of Chapter 4 in the thesis structure

4.1 Introduction

This chapter documents the exploration of smart card data as a means for measuring customer fluctuation within public transport markets. The use of smart card data as a measurement approach has been selected as it provides a large population sample. This smart card sample tracks individual smart card IDs and the associated travel behaviour over time. This might allow for the capture of annual changes in travel behaviour, as well as the ability to review variations at different time scales and analysis periods. In Melbourne, all public transport users are required to have an active smart card to travel on public transport legally and as such it provides a key resource for measuring user behaviour. A rule-based segmentation method based on longitudinal travel patterns is developed to identify different types of customers (new, retain, return and lost) using smart card data. The analysis focuses on measuring customer fluctuation rates and identifying the limitations and challenges for future studies.

This chapter begins by stating the aims of the smart card analysis, followed by a brief description of the research approach. The results are presented, including the measurement of customer fluctuation and the impacts of variations in the temporal unit of analysis (e.g., monthly and weekly travel patterns). A discussion of the customer fluctuation results and limitations of the approach and challenges follows. This includes an analysis of total travel volume and customer fluctuation segments and a more detailed review of the new and lost user segments and the variability within these segments.

Finally, this chapter concludes with an identification of the next steps to be taken to improve the measurement approach for customer fluctuation.

4.2 Aims

The secondary data analysis undertaken in this chapter focuses on providing the first attempt at answering Research Aim I:

> I. To develop, measure and apply a new concept for market change analysis based on change segments.

More specifically, this chapter seeks to meet the following research objective:

> RO3: To explore approaches to measure market change segments

The research approach for this initial application of the concept of customer fluctuation to public transport ridership data is outlined in the following section.

4.3 Research Approach

An overview of the research approach, highlighting how the method answers the overarching research question and research sub-questions, is shown in Figure 4.2.

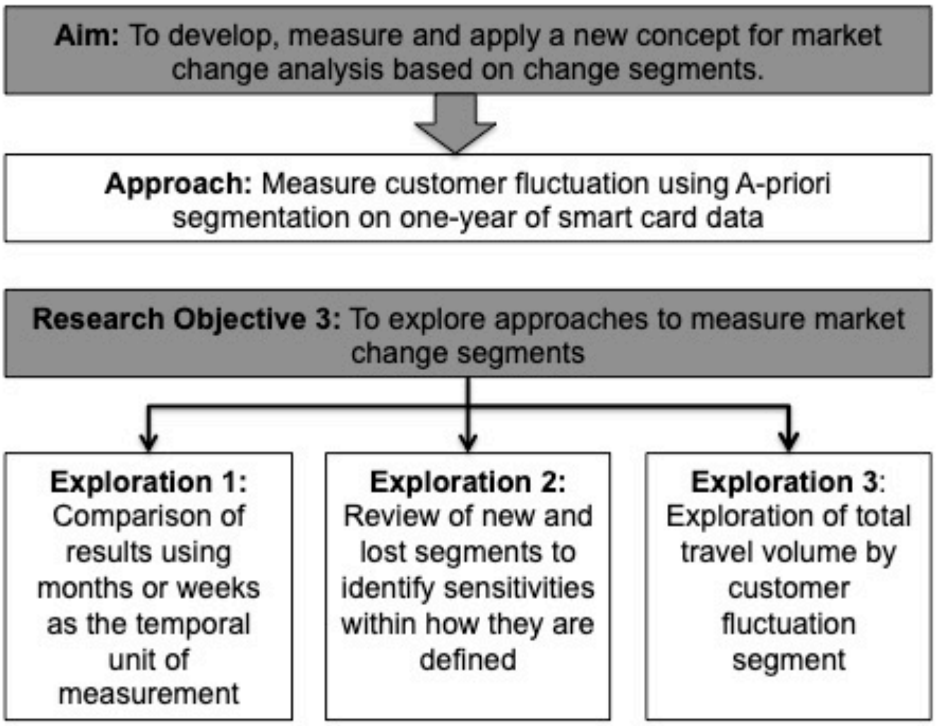

Figure 4.2 - Overview of Research Approach used in Chapter 4: Measuring Customer Fluctuation using Secondary Data

This initial measurement approach for market change analysis adopts an a-priori segmentation approach. A-priori segmentation uses a predetermined target variable (dependent variable) to aid in meeting the research aims. This target variable, in effect, is a predefined segment (based on: Elmore-Yalch, 1998, Kieu et al., 2015b). Using this approach, clusters are defined by pre-selected characteristics which reflect the groups outlined by the concept of 'customer fluctuation': new, lost, retained and returning users. The customer fluctuation concept for measuring market change is discussed in detail in Chapter 3.

The challenges for a-priori segmentation are the determination of the appropriate measurement period and temporal units of analysis of each group. This analysis seeks to explore different definitions and approaches in the literature to measure changes in an individual's ridership patterns. The initial analysis utilises one year of smart card data, collected between October 2016 and

September 2017. One year is used as the measurement period and the temporal unit of analysis includes both weekly and monthly approaches.

The data is provided by Transport for Victoria. It consists of travel details for a random sample of smart card IDs. In Victoria, the smart card system is called ‘myki’, and this term will be used interchangeably with the smart card within this section. This system is used on all public transport modes in Melbourne (bus, train and tram). The information associated with each entry includes the unique myki ID, the date of the trip, the mode used, the stop location and geographical coordinates. All myki data has been provided anonymously to protect the privacy of myki users.

As discussed in Chapter 3, each smart card was then segmented based on the temporal characteristics of their ridership identified in Table 4.1. The temporal unit is provided in months; for the weeks-based analysis, one month is considered equal to four weeks. This time frame is kept roughly equivalent so that we can measure the differences in sensitivity to change between the two approaches. For example, a user is considered lost if they miss more than two consecutive months’ travel, by using months they can travel at any point during the third month to not be identified as lost. When using weeks as the measurement unit, they must travel within the next week, or be identified as lost.

Table 4.1 - A-Priori Segmentation Rules for Measuring Customer Fluctuation with Coding Parameters for Months and Weeks Based Analysis

Segment	Description	Coding Parameters
New Users	Where the first active trip occurs after December '16 and there are continuous trips each month with no more than one month 's/four weeks break the user classifies as a **'new user**.'	After the first 2 months
New Users	Where a smart card/user appears in the last two months of the study period (September or October), they classify as **'new users'**. This rule includes users that appeared in September but not October, as they did not yet qualify as 'lost users’ due to a lack of time for consecutive misses.	Last 2 months of the measurement period
Lost User	Those that start travel and then miss two or more consecutive months and do not return in the measurement period classify as a **'lost user.'**	Miss more than 2 consecutive months and do not return
Retained user	Where users have a minimum of one active trip per month for the entire period, with no more than two consecutive months break the user classifies as a **'retained user**’. These breaks are to allow for travel/annual leave	One active trip per month, no more than 2 consecutive months’ break
Retained user	Smart cards/users were still identified as **'retained users**’ if they had multiple breaks in travel if these breaks did not occur in consecutive months and accounted for less than 25% of the overall period.	Missed less than 3 months’ overall
Returning User	Smart cards/users that had returned once were assumed likely to return and thus were not categorised as lost if they had two consecutive months of no travel after travelling and returning to travel within the measurement period. These users remained classified as a **'returning user'.**	n/a
Returning User	If there is greater than 25% of the measurement period without travel (consecutive or non-consecutive) followed by travel, the smart card/user will be a "**returning user”**.	Missed more than 3 months’ overall
Returning User	Those that have an irregular travel pattern inconsistent with the above definitions are known as a **'returning user'**.	n/a

Several assumptions have been utilised for the development of the a-priori segmentation rules described above. This included assumptions that attempt to find the appropriate thresholds for each behaviour change segment (new, lost, retained, returning). These thresholds were created to allow for users to undertake standard variations in travel over the course of the year (such as short holidays) without necessarily influencing their ridership segment classification. Further exploration of these thresholds, for the new and lost segments, and the impact on measurement is reviewed later in this chapter. These segments were targeted for additional review as they are impacted by both the measurement period and the temporal unit of measurement.

For further details on the assumptions associated with these segmentation rules, please refer to Chapter 3: Concept Development and Research Approach.

The results of this analysis are detailed in the next section.

4.4 Results

The following section provides the results from the initial exploration of Customer Fluctuation using one year of smart card data.

4.4.1 Impact of Temporal Unit of Analysis

Monthly Unit Based Analysis

Table 4.2 customer fluctuation results using **months** as the temporal unit of analysis.

Table 4.2 –Customer Fluctuation Results Using Months as the temporal unit of analysis

	Train		Bus		Tram		Total Sample[1]	
	No. of smart cards	% of train sample	No. of smart cards	% of bus sample	No. of smart cards	% of tram sample	No. of smart cards	% of tram sample
New	14,475	21.9%	7,087	23.1%	10,166	20.4%	17,770	23.6%
Lost	30,721	46.4%	15,876	51.8%	26,611	53.5%	31,881	42.4%
Retained	4,407	6.7%	1,641	5.4%	2,061	4.1%	6,438	8.6%
Returning	16,577	25.0%	6,064	19.8%	10,888	21.9%	19,139	25.4%
Total	**66,180**	**100%**	**30,668**	**100%**	**49,726**	**100%**	**75,228**	**100%**

Notes: All figures are based on the smart card ID Sample between October 2016 and September 2017. For this assessment, each smart card ID is assumed equivalent to one user.
[1] The total sample provides a review of the entire sample of smart cards provided regardless of the mode chosen. This means that both multi-modal riders and single modal riders are accounted for. This explains why the total sample is not equal to the sum of each mode-based sample.

The findings illustrate that:

- The lost user segment dominates all smart cards by mode (46-54%) followed by returning market segments (20%-25%) and new riders (20-23%). The smallest category across all mode is the retained riders (4-7%).
- The train smart card market has the lowest rate of lost rider share at less than half of the total sample of train users (46%). It also has the highest rates of retained riders (7%) and returning ridership (25%). The train falls between the bus (highest) and tram (lowest) in terms of the proportion of new users.
- Over half of the tram users are lost during the year. They also have the lowest rate of customer retention (4%).
- The bus is somewhere between the train and tram in terms of customer fluctuation segment share.
- The total sample[1] are in general consistent with the individual mode results, though the share of the lost user segment is lower than for the mode results.

A Pearson Chi-Square test was run to identify whether there is a statistical difference between the size of segments for each mode. Chi-square tests indicated that there was a statistically significant relationship between the customer fluctuation segments and each mode, χ^2 (6) =948.689, p<.000. The Cramer's V (0.057) identified that the effect of mode on the customer fluctuation segment was small.

Weekly Unit Based Analysis

An identical assessment was also conducted using **weeks** as the period of temporal analysis. The results of this segmentation are in Table 4.3.

Table 4.3 - Customer Fluctuation Results Using Weeks as the temporal unit of analysis

	Train		Bus		Tram		Total Sample[1]	
	No. of users	% of pop.	No. of users	% of pop.	No. of users	% of pop.	No. of users	% of pop.
New	11,477	17.2%	5,702	18.5%	7,942	15.9%	10903	13.6%
Lost	28,293	42.5%	14,927	48.5%	25,034	50.0%	32918	41.2%
Retained	2,327	3.5%	843	2.7%	833	1.7%	3044	3.8%
Returning	24,549	36.8%	9,323	30.3%	16,272	32.5%	33049	41.4%
Total	**66,646**	**100%**	**30,795**	**100%**	**50,081**	**100%**	**79914**	**100%**

Notes: All figures are based on the smart card ID Sample between October 2016 and September 2017. For this assessment, it is assumed that each smart card ID is equivalent to one user. Percentages show the proportion of each segment within the mode sample.
[1] The total sample provides a review of the entire sample of smart cards provided regardless of the mode chosen. This use means that both multi-modal riders and single modal riders are included. The inclusion of multi- and single-mode travel is why the total sample is not equal to the sum of each mode based sample.

[1]The total sample includes multi-modal trips and is not a total of the other results. Hence the share of the total sample segments is different from the total shares for all three modes

The findings of the weekly segmentation are broadly similar to the results of the monthly analysis. Specifically, the customer fluctuation segmentation using weeks provides the following results:

- The lost rider segment dominates all smart cards by mode (42 - 49%) followed by the returning users (30 - 37%) and new riders (15 - 19%). The smallest category across all mode is the retained riders (4 - 7%).
- The train smart card market has the lowest rate of lost rider share (43%). It also has the highest rates of retained riders (4%) and returning ridership (37%).
- Tram markets smart cards lose half their riders during the year and have the lowest rate of new ridership (16%). This low rate of rider acquisition is coupled with the lowest rates of retained rider smart cards (2%). Returning ridership share is also the lowest for the tram (33%).
- The bus is somewhere between the train and tram in terms of market fluctuation market shares of smart cards. However, the bus had the highest rate of new users (19%), coupled with a higher rate of lost users (49%) and the lowest rate of returning users (30%).

Chi-square tests indicated that there was a statistically significant relationship between the customer fluctuation segments and each mode, $\chi2$ (6) =1142.643, p<.000. As with the monthly analysis, the Cramer's V (0.062) shows that the effect of mode on the customer fluctuation segment is small.

To provide a more detailed analysis, Table 4.4 presents a comparison of the customer fluctuation segmentation results when using a weekly or monthly temporal analysis unit.

Table 4.4 – Comparison of Customer Fluctuation Segments using weeks or months as temporal analysis unit

	Train		Bus		Tram		Total Sample	
	Months	**Weeks**	**Months**	**Weeks**	**Months**	**Weeks**	**Months**	**Weeks**
New	21.9%	17.2%	23.1%	18.5%	20.4%	15.9%	23.6%	13.6%
Lost	46.4%	42.5%	51.8%	48.5%	53.5%	50.0%	42.4%	41.2%
Retained	6.7%	3.5%	5.4%	2.7%	4.1%	1.7%	8.6%	3.8%
Returning	25.0%	36.8%	19.8%	30.3%	21.9%	32.5%	25.4%	41.4%
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

The comparison between temporal units shows that:

- The returning customer segment was the only segment that has a higher share of smart cards/ users between a month- to week-based assessment, with all other segments having a lower segment share when using weeks rather than months.
- Both assessments found that train markets had the largest share of retained users and tram markets had the lowest share. The inverse was true for the share of lost users. The tram has the highest share of lost users, and the train has the lowest share.

- The bus had the highest share of new riders, though it was still roughly equal to the rates for both the train and tram.
- The increased sensitivity when using weeks as the temporal unit of analysis made it harder to identify continued travel patterns and increased the proportion of users that were identified as returning, with an irregular pattern of travel.

The temporal unit tests highlighted that the use of months and weeks produced largely similar results, with the most substantial difference the increase in the returning user segment when using weeks. Though it is not possible to assess the relative accuracy of either measurement, the increase in returning users and decrease in the proportion of all other segments might indicate that the use of weeks is too sensitive as a temporal unit of measurement for this purpose. Further, as the returning user segment accounts for all patterns inconsistent with new, lost or retained use; this may limit the ability to derive clear conclusions about individual travel patterns. It is also considered that unlike the studies conducted by Mason et al. (2011), this work seeks to measure the entire market rather than separating commuter and leisure users. This is to provide a truer picture of market function. Though the variation is small, it is suggested that months are sufficiently sensitive to allow for both high-frequency users (e.g. commuters) and lower frequency users (e.g. for leisure and other travel) to be appropriately segmented. As such, weeks may be a more appropriate temporal unit when investigating only those that commute or otherwise travel regularly (several times per week) on public transport.

As the differences between the two temporal units were not substantial, further consideration is given to the practical realities of measurement. The analysis of 12 monthly units rather than 52 weekly units allows for the development of a streamlined approach. Further, Cramer’s V results provided evidence that the factors that influence customer fluctuation segment may be consistent across the three modes studied, as there are no substantial differences in the size of customer fluctuation segments between modes. Primary data sources are required to identify the reasons influencing customer fluctuation behaviour. Due to the timing of this thesis, this is being completed through a retrospective survey, which relies on participants to recall their past travel behaviour. It is unreasonable to expect that participants would remember their travel for the past year at the scale necessary to complete a weekly analysis. As such, a monthly scale would be more appropriate and maintaining that scale here would allow for direct comparisons on the application of customer fluctuation.

It is for these reasons that a month based temporal unit of measurement is selected as the preferred approach for the remainder of this thesis.

As this is the first exploration of the customer fluctuation concept, the temporal unit is not the only element that requires further exploration. It is also useful to operators to know the ‘purchasing behaviour’ of each segment – that is, the average number of trips each segment makes. This is briefly explored in the next section as to how the scale of each segment might provide additional insights.

4.4.2 Total Travel Volume Analysis by Customer Fluctuation Segment

Table 4.5 shows the size of each market segment in Melbourne by comparing the number of smart cards to the total number of trips. The smart card data provided for this research also provided information on the number of trips taken by individual smart cards. This information was aggregated for each month and as such the number of trips for each segment could be calculated. The segmentation results based on smart cards could then be compared to the proportion of trips taken by each segment group.

Table 4.5 - Comparison of the proportion of smart cards in each Customer Fluctuation segment to the proportion of total trips

	Train		Bus		Tram		Total Sample	
	% of smart cards[1]	% of total train trips[2]	% of smart cards[1]	% of total bus trips[2]	% of smart cards[1]	% of total tram trips[2]	% of smart cards[1]	% of total trips[2]
New	21.9%	23.2%	23.1%	26.8%	20.4%	23.8%	23.6%	18.3%
Lost	46.4%	22.0%	51.8%	26.5%	53.5%	27.0%	42.4%	19.6%
Retained	6.7%	39.1%	5.4%	32.1%	4.1%	30.0%	8.6%	27.3%
Returning	25.0%	15.7%	19.8%	14.6%	21.9%	19.2%	25.4%	34.8%

Note: [1] It is assumed for this assessment that one smart card is equivalent to one user
[2] A trip is recorded as one touch on and may exclude instances of continuation trips.

The findings indicate that:

- Although lost users are the largest group of card types (46 - 54%), they represent a proportionally smaller share of annual trips made (22% - 27%). This lower share of trips is further investigated in the following section.
- Although the retained user segment is the smallest (4-9% of cards), they represent the largest segment of total annual travel for the train (39%), bus (32%), tram (30%) and the second-largest for the total sample (27%). This finding is consistent with the marketing literature on retaining existing customers rather than attracting new users for growing customer markets, as well as general knowledge in the public transport industry (Kieu et al., 2015b, Reinartz et al., 2005).

- New users are the second-largest share of total travel for all modes (23 – 27%), with the highest percent of trips from new bus users (27%). However, new users have the lowest share for overall card use.
- Returning users were the smallest percentage of total trips for all modes (16% for train markets, 15% for bus markets and 19% for tram markets) except for multi-modal use across the sample, where returning users were responsible for the highest volume of trips (35%).

This section explored trip volume as a potential alternative approach for the measurement and understanding of market change segments. Overall, this found that trip volume is an indicator of market segment purchasing behaviour **and would be a beneficial area for further analysis following the completion of a final measurement tool.** This consideration is an important consideration of the practicalities of measurement and what it might mean for public transport operators.

The next section explores different possible definitions of new and lost users, which are both sensitive to the length and timing of the measurement period as well as the temporal unit of measurement.

4.4.3 Exploring Variability within 'Lost' Smart Card Patterns

Lost users dominated all public transport markets explored (train, bus and tram), yet this category is sensitive to variations in segmentation rules and definitions. To better understand this group, Table 4.6 shows a breakdown of lost riders divided into three new sub-categories:

- 'One-off' smart cards, that rode for only one month within the study period, excluding those that only travelled in the first month of study;
- 'Seasonal' riders who used public transport for two or more months but less than six months in the middle or end of the time studied; and
- ‘Retained, then lost’ riders, who were using public transport in the first months of study, in a pattern like retained users, and then no further trips were made at all during the period of analysis.

These patterns are also illustrated in Figure 4.3.

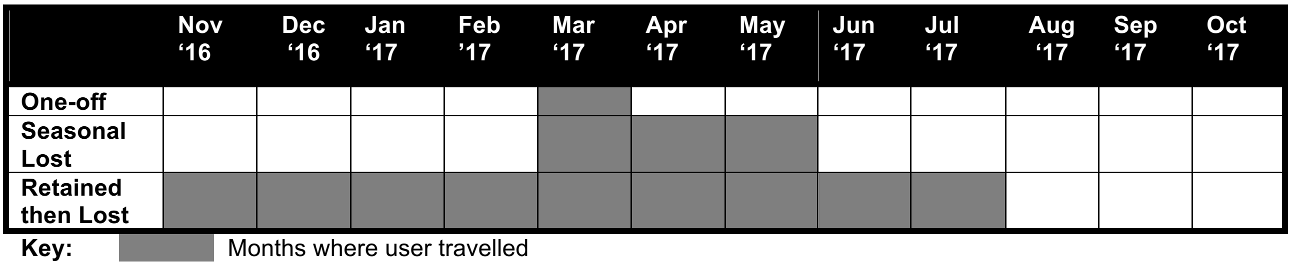

Figure 4.3 – Typical pattern for each lost ridership category

Table 4.6 – Investigation of Differences Between 'lost' rider Smartcard Segments

	Train			Bus			Tram		
	No. of lost cards	% of lost users	% of train sample	No. of lost cards	% of lost users.	% of bus sample	No. of lost cards	% of lost users	% of tram sample
One-off	16,859	55.1%	25.5%	8,523	53.8%	27.8%	16,862	63.6%	33.9%
Seasonal	8,779	28.7%	13.3%	4,327	27.3%	14.1%	6,089	23.0%	12.2%
Retained, then Lost	4,955	16.2%	7.5%	2,986	18.9%	9.7%	3,574	13.5%	7.2%
Total	**30,601**	**100.0%**	**46.2%**	**15,836**	**100.0%**	**51.6%**	**26,525**	**100.0%**	**53.3%**

This analysis indicates that:

- A significant portion of lost user smart cards are 'limited users' who catch a mode for no more than one month of the year. These limited users account for almost two-thirds of lost tram users (64%) and slightly over half the users for the bus (58%) and train (55%). One possible explanation is that smart card data captures tourists visiting Melbourne, as well as a proportion of users that have lost smart cards, or have a household smart card that is used irregularly.
- This group is followed by a significant proportion of lost smart cards that appear seasonally throughout the year. This accounted for approximately 28% of lost bus and train riders and 23% of lost tram riders. The peak seasons for seasonal travel differed by mode. For the train, most seasonal trips were witnessed between December and June. For bus and tram users the seasonal peak was between January and May.
- 'Retained then lost' customer smart cards (who were travelling for a minimum of two consecutive months from the start of the study period and then did not travel again) accounted for only 16% of lost train users, 13% of bus users and 14% of tram users.

This analysis suggests that current definitions are capturing different sub-segments of lost users. If customer fluctuation is adopted as a method of measurement for public transport markets, identifying the proportion of 'retained, then lost' users may be a valuable secondary step.

4.4.4 Exploring Variability within 'New' Smart Card Patterns

The proportion of new user smart cards is also of interest due to the low numbers of the total population that are "new" to public transport over the period. There was not a considerable variance amongst the proportion of new rider smart cards attracted to each mode. The tram had the highest share of new rider smart cards (13%), closely followed by the bus (11%) with the train relatively close behind (9%).

Based on the a-priori rules, many users that meet the criteria for a new user appeared only in the last two months of the study period. This finding accounts for 59% of new train users, 70% of bus users and 79% of new tram users. This is an issue as these users may exit the market in the months after the measurement period. If the measurement period were extended, many of these users would be reclassified as returning or lost. **This indicates that there may be a need for an extended measurement period, including a buffer, to verify customer fluctuation classifications**.

These results highlight an important issue in identifying new riders using the a-priori segmentation rules; new riders are only classified as new if they continue to use public transport over the (annual) study period i.e. they are both new and continuous. This approach was adopted to ensure clearly defined categories that are independent of one another. However, this does imply that some new riders will become lost, and others may then return late in the year. In each case a rider is both new and lost and a returning user depending on when they were measured.

This raises the question; 'what is a new rider?'. Our current assumption is that new riders are those that enter the market and proceed to ride continuously without a break. This may not be a reasonable expectation. Table 4.7 aims to explore this further; it shows the proportion of 'new' rider smart cards that travelled continuously but also those market entries that were classified as lost or returning.

Table 4.7 – Total Number of New Rider Smart Card Including Market Entries That Have Become Lost or Returning

	Train			Bus			Tram		
	No. of smart cards	**% of total new**	**% of total train sample**	**No. of smart cards**	**% of total new**	**% of total bus sample**	**No. of smart cards**	**% of total new**	**% of total tram sample**
Market entry and retained	14,475	31.9%	21.9%	7,103	31.3%	23.2%	10,191	28.3%	20.5%
Market entry and then lost	23,549	51.9%	35.6%	12,997	57.3%	42.4%	20,862	57.9%	42.0%
Market entry and then returning	7,347	16.2%	11.1%	2,598	11.4%	8.5%	4,986	13.8%	10.0%
Total	**45,371**	**100.0%**	**68.6%**	**22,698**	**100.0%**	**74.0%**	**36,039**	**100.0%**	**72.5%**
Note: New users are identified as any individual card that had the first appearance after the initial two months of the study period. As new users can be identified as lost or returning based on the final date of assessment, this exploration has included all those that commenced travel after the first two months of the assessment period.									

Table 4.7, shows that there is a significant difference in 'new users' if we include those who stop travel or return. In total new user smart cards including all variations would represent 74% of bus users but only 23.2% if only the retained new users are included. Similar patterns are evident for other modes. **Clearly, we may conclude that the definition of what are new and lost user smart cards depends on the period over which the analysis is undertaken and the time segments were chosen.**

4.5 Discussion

The Measurement Approach

This chapter has completed the first application of customer fluctuation to public transport market data. The measurement approach has produced results that are logically consistent with our current knowledge of public transport markets. The results suggested that all transport modes in Melbourne public transport markets were dominated by lost smart cards/users, accounting for nearly half of all smart cards analysed. Returning smart cards (20 – 25%) and new smart cards (20 – 23%) were the next largest groups. The smallest segment across all modes were retained riders (4 – 7%). A chi-square test found the size of market segments differed between modes. However, a Cramer's V test identified that the correlation (or strength) of the association between the user and customer fluctuation segment is quite small, suggesting that the differences between modes are not large.

An assessment comparing weeks and months as the temporal unit of analysis was also undertaken. This analysis suggested that month was the preferred temporal unit of measurement for the data. Although both temporal units were suitable, months had the advantages of being less sensitive to short-term variations in travel behaviour.

Regardless, the initial application of customer fluctuation found there are some limitations with Melbourne's smart card data that must be considered when interpreting these results. Melbourne's smart card data does offer benefits as all public transport riders within Melbourne are required to hold a smart card to travel legally, as paper ticket options do not exist. Although smart cards can collect large volumes of data, the system has not been designed with the use of this data in mind thus presents several limitations, including:

- Smart cards cannot be equated to individual riders as users may hold multiple smart cards or lose and need new smart cards regularly. Conversely, people can share one smart card with friends or family members.
- Smart cards expire after four years of purchase, requiring a new card with a new ID. These IDs are not linked between new and old cards.
- Smart card data cannot perfectly capture the true number of riders due to the prevalence of fare evasion, though the frequency of evasion from card holders is unknown.
- Smart card data is anonymised and individuals cannot be directly linked to accurate demographic data. This can be partially addressed through modelling; however, it still relies on demographic assumptions based on location. This limits the ability to understand any demographic differences within the customer fluctuation segments measured.

These limitations are thought to have the largest impact on the proportion of lost riders (increase) and retained riders (decreased), when measuring customer fluctuation. Several potential strategies to measure or reduce the impact of these limitations on results are discussed in Chapter 5.

Exploration and Opportunities for Refinement

Research objective 3 is to measure customer fluctuation and to explore measurement approaches for this new concept. To fulfil this objective, we conducted a review of the new and lost user segments. These segments were identified as particularly sensitive to how they were defined, due to their dependence on what measurement period is selected and the month in which it starts and ends.

A high share of new users (50% of new train users, 60% of new tram users and 23% of new bus users) were those that appeared in the last two months of the measurement period. As such these were identified as new users, though their travel patterns were not long enough to confirm the appropriate segment. This identifies an issue with the current measurement approach; new users are required to travel consistently (like a retained user) once appearing after the first two months of the measurement period. This is partially to account for the inability to tell whether users are new to the service, or whether they are a returning user and this is their first appearance within the one-year measurement period. To improve the measurement of new users a longer measurement period should be explored. There may also be value in exploring further restrictions at the start of the measurement period, such as a buffer to confirm new use.

There was a similar pattern with lost riders that were identified as lost because they did not have a chance to return before the end of the analysis period. This sensitivity to timing is a limitation of the existing approach; research should investigate the inclusion of a buffer period at the start and end of the analysis period. A buffer allows for behaviour to be confirmed by a few additional months of travel to allow for these issues and their impact on the final segmentation results.

Overall, the limitations of a 12 month measurement period were highlighted in this analysis, suggesting that a longer period would provide greater clarity of results.

4.6 Conclusion and Next Steps

This chapter applied the concept of Customer Fluctuation to public transport markets and developed the measurement approach using smart card data. This initial assessment identified that the rates of customer fluctuation were similar across all modes of public transport, with lost users the biggest proportion of the population, followed by new and returning users which shared similar proportions, and a small proportion of retained users.

Although initial results are promising, there are opportunities to further refine the measurement approach and the data used. This analysis found that calculating customer fluctuation using a one-year timeframe required several assumptions when categorising users based on travel occurring at the beginning or end of the timeframe. **An alternative approach is proposed where two years of data is utilised to allow for a confirmation 'buffer' at the beginning and the end of the measurement period to allow for the ambiguous travel patterns of lost and new users to be better measured**. Chapter 5 will complete a detailed analysis of a smaller sample of smart card data that covers two years of travel. This analysis will investigate the use of a buffer to confirm whether riders have been appropriately segmented.

Chapter 5: Measuring Customer Fluctuation using Secondary Data, Part Two

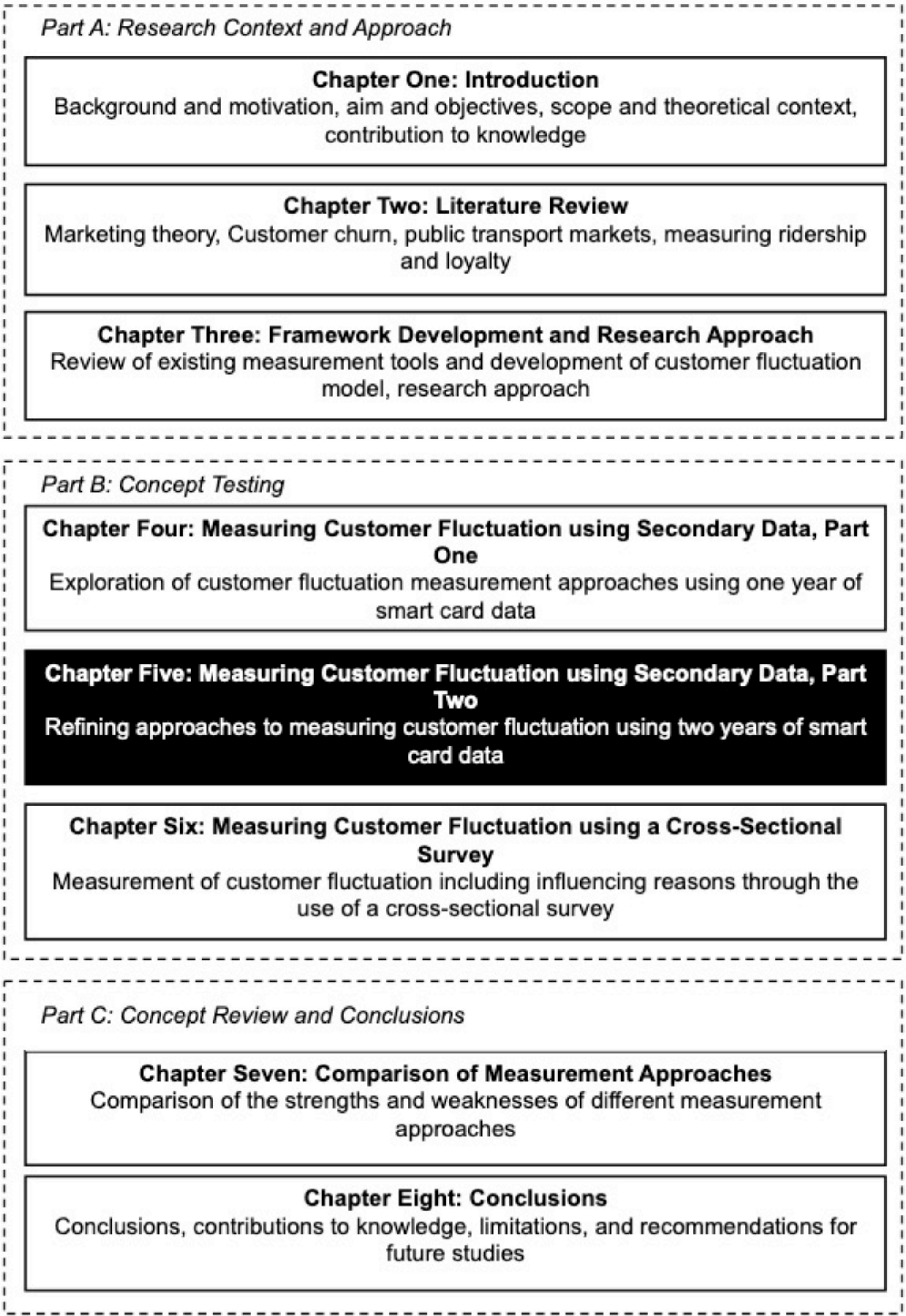

Figure 5.1 - Position of Chapter 5 in thesis structure

5.1 Introduction

This Chapter is the second chapter that focuses on measuring customer fluctuation using secondary data; building on the findings of Chapter 4 (Part One). The previous chapter documented the exploration of an approach to measure customer fluctuation using one year of smart card data. The findings of this analysis produced a preliminary measurement of customer fluctuation segments and some early insights. However, it was clear that there were several areas where the measurement approach could be refined. Findings suggested the need to extend the measurement period and create an analysis buffer to more appropriately capture customer fluctuation rates. Questions were also raised about the impact of lost or expired smart cards on the segmentation results and the lack of socio-demographic data to aid in understanding patterns of fluctuation in travel behaviour.

This chapter uses a new data set that contains a random sample of approximately 4,000 smart cards for each mode over a period of two years (October 2016 to September 2018). The data set includes additional details compared to the data set adopted in Chapter 4, including card type and the card activation date. These details allow for the additional exploration of demographic comparisons (using fare type codes) and the impact of card expiry on fluctuation rates. It is noted that due to these additional details and the volume of data generated over a two year period, the sample size is much smaller for this study than the one year data set used in Chapter 4. Regardless, the same a-priori segmentation approach is adopted for this chapter, using months as the temporal unit of assessment. Any other adjustments to the approach have been minor and will be discussed further in this chapter.

This chapter begins by stating the aims of the secondary data analysis utilising two years of smart card data. This is followed by a review of the adjusted research approach building on the approach in the previous chapter. The results are then presented, including the following analyses:

- Customer fluctuation segment sizes using two years of data, with and without a 'verification buffer'
- A comparison of one- and two-year segmentation results
- Customer fluctuation by card type
- Trip volume by customer fluctuation segment
- An in-depth exploration of new and lost rider segments based on card initialization and expiry

These results are followed by a discussion section that outlines the key findings of the overall smart card data assessment, including the identification of limitations in these methods. The final section outlines the key conclusions from this analysis.

5.2 Aims

The secondary data analysis undertaken in this chapter focuses on providing a refined approach to assist in meeting Research Aim I. This aim is re-iterated below for reference:

> I. To develop, measure and apply a new concept for market change analysis based on market change segments.

More specifically, this chapter seeks to meet the following research objective:

> RO3: To explore approaches to measure market change segments

The adjusted research approach developing on the findings of the previous chapter is outlined in the following section.

5.3 Adjusted Research Approach

Chapter 4 provided results from the first application of customer fluctuation to public transport markets using smart card data. These results identified a statistically significant relationship between mode and customer fluctuation segments. However, several limitations were also identified with the initial research approach.

The analysis found that calculating customer fluctuation using a one-year timeframe required several assumptions when categorizing card users. The one-year length of the measurement period did not account for seasonal variations in travel, for example, known peaks due to school holiday periods. These assumptions impacted on the segmentation of new and lost users. As well as reflecting seasonality, the measurement period may result in returning users being identified as new when appearing later in the period, and lost users would be better categorised as returning but had not returned by the end of the measurement period.

To address the impacts of the measurement period, a two-pronged revision to the approach has been adopted. This approach extends the time frame to a total of <u>two years</u> and introduces the use of a '<u>verification buffer</u>'. The verification buffer adds a three-month 'buffer' period at the beginning and end of the measurement period. These 3-month periods are used to validate the results from the active segmentation window taken over 18 months, by identifying users that may have been segmented incorrectly dependent on when they appear in the measurement window

The second limitation identified in Chapter 4 is that it is difficult to link smartcard data with demographic information. One approach used in the literature is to attempt to derive a home location based on smart card usage data and link the average neighbourhood socio-demographic characteristics to the user. However this process is time consuming and still relies heavily on many assumptions, such as ascribing the average socio-demographic characteristics of an area to an individual (Ma et al., 2013).

A preferred approach is to derive available demographic data from the smart cards themselves, where possible. The two-year data sample provided to the research team includes card type, which is a potential avenue for identifying a limited set of socio-demographic characteristics. The ticketing and card types within Victoria are not directly linked to demographics but they do imply certain demographic features, and, for ease of analysis, they were separated into the following groups:

- 'Full fare' includes cards identified as 'full fare' and 'default full fare';
- 'Child' includes cards identified as 'Child Concession 4 yrs to <= 16yrs' and Child Concession 5 yrs to <= 18yrs';
- 'Student' includes any cards identified as 'Tertiary Student Concession', 'Secondary Student Concession' or an 'International Student';
- 'Seniors' includes any cards identified as 'Pensioner Concession Card Holder PC', 'Retired Employee Dependent Travel Pass' and 'Retired Employee Travel Pass';
- 'Other' concessions include all other concessions or reduced fare card types.

The inclusion of these card type groups allows for the exploration of a limited set of demographic features of different customer fluctuation segments. This also provides an opportunity to measure trip volume against card type to identify and further patterns of travel.

A final adjustment explored in this chapter is the inclusion of smart card activation and card expiration dates. This allows for further exploratory data analysis into how lost, expired, or new smartcards impact customer fluctuation segments.

No other elements of the research approach adopted in Chapter 4 have been changed to allow for consistency of measurement and comparisons between data sets. As such, all a-priori segmentation rules and categorisations are retained. This section has continued to use months as the period of temporal analysis as there was only a moderate statistical difference when utilising weeks. It is noted that the two-year data analysed for this sample is comprised of a smaller sample of smart cards than the one-year data set, with approximately 4,000 cards sampled for each mode. This has been required due to the increased volume of trips recorded over the longer period, which creates an exponentially larger data set over time. Regardless, the sample is still of a sufficient size to provide statistically significant results for metro Melbourne.

5.3.1 The Verification Buffer

The purpose of the verification buffer is to improve accuracy when identifying customer fluctuation segments at the start and end of the measurement period. This aims to mitigate the impact of when the measurement period is placed (start and end dates) on the numbers of new and lost users at the end of the measurement period and, to a lesser extent, new users at the start of the measurement period. These are users that are identified within a category due to their short length of time appearing in the study, where it is reasonable to believe that if we continued the length of the measurement period, they may demonstrate different behaviours and belong to different segments. Similar adjustments were made by Chu (2015), however, in Chu's work, cards that were only partially active within the two year analysis period were excluded for assessment.

An example of how the verification buffer is applied is illustrated in Figure 5.2. This shows how the buffer is split with 3 months provided at the beginning and 3 months at the end of the measurement period for verification. The 3 months at the beginning is applied to more accurately identify new users, the 18 months is used to segment users, and the last 3 months is to confirm whether that segmentation result was correct.

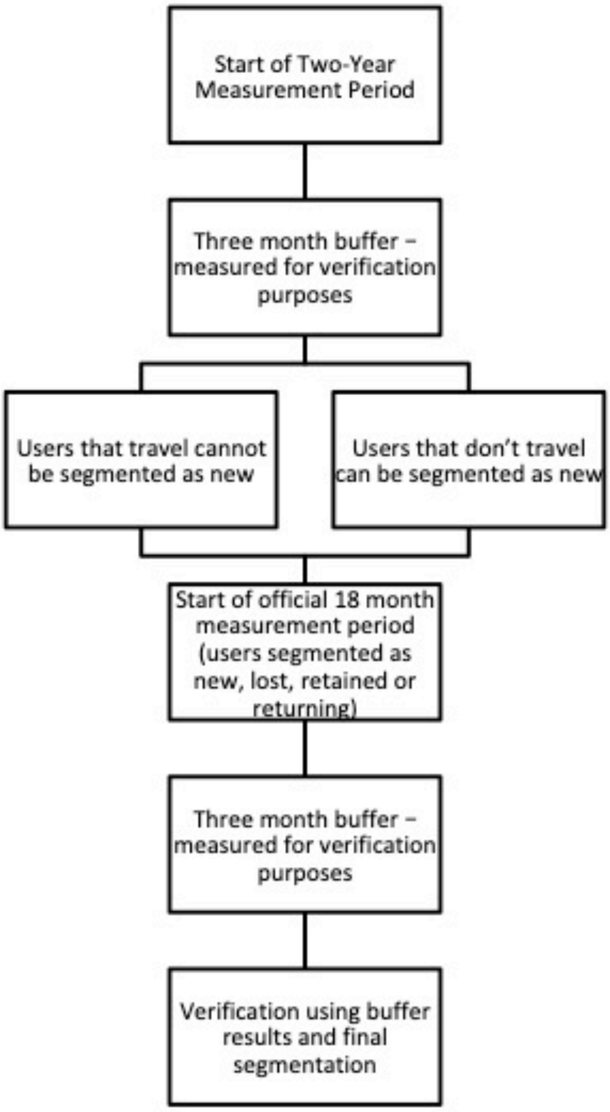

Figure 5.2 - Process Diagram for the application of verification buffer

The application of the verification buffer requires the completion of the a-priori segmentation approach for both two years and 18 months (excluding the first three months and last three months of the complete two year data set). The segmentation for 18 months is then manually reviewed for accuracy using the data from the verification buffer (the first and last 3 months of the two year segmentation). New users are confirmed by identifying whether they travelled in the 3 month verification buffer, and thus should be segmented as retained or returning users. The verification buffer at the end of the measurement period is then used to confirm all segments. This simple process ensures that users have not been segmented based only on their behaviour at the end of the measurement period, but rather their ongoing travel patterns.

This analysis is conducted to test the effectiveness of a verification buffer. The inclusion of a verification buffer should ensure that seasonal peaks and lulls, such as school times or holiday periods, do not have a significant impact the segmentation results.

5.4 Results

Market segmentation results are compared using three different approaches: two years of data, 18 months with a buffer and the one-year results from Chapter 4. The same a-priori segmentation rules used in Chapter 4 are adopted. Further analysis is conducted, including a travel volume analysis and segments by card type. All three applicable modes of public transport within the Melbourne context (train, bus, and tram) were analysed.

5.4.1 Comparison of Results with and without Verification Buffer

Table 5.1 provides the results of the a-priori segmentation for the full two years of data.

Table 5.1 - User Breakdown by Mode into Customer Fluctuation Segment over 2 years, WITHOUT buffer

	Train		Bus		Tram	
	No. of users[1]	% of users	No. of users[1]	% of users	No. of users[1]	% of users
New	107	2.7%	100	2.7%	82	2.4%
Lost	1,725	43.1%	2,030	54.7%	1,793	52.9%
Retained	87	2.2%	46	1.2%	49	1.4%
Returning	2,081	52.0%	1,537	41.4%	1,466	43.2%
Total	4,000	100%	3,713	100%	3,390	100%

Note: All figures are based on a two-year period between October 2016 and September 2018.
All modes had a randomly selected sample of approximately 4000 smart cards.
[1]Users have been defined as an individual smart card ID

The findings of this segmentation using the existing a-priori segmentation rules and two years of smart card data identifies that:

- The lost rider segment dominates for both the bus and tram at roughly 50% of the card population. This was followed by returning users (41%-52%). The smallest segment of the population was retained users (both approximately 1-2%), which was only slightly smaller than the proportion of new users (about 2%).
- For the train, returning users dominate (52%) with lost users the second largest user segment (43%). Retained users were the smallest proportion of train users (2%), and new users were again only slightly larger (3%). The train had the highest proportion of retained users of all modes.

This analysis indicates new and retained user rates were relatively consistent across all modes, however, returning and lost users show the greatest variability between the train and the tram/bus. Tram and bus have a majority of lost users (55% and 53%, respectively). These findings might provide a preliminary indication that Melbourne train markets exhibit lower rates of customer fluctuation, due to the dominant returning share and lower rates of lost users. A Pearson Chi-Square test was run to identify whether there is a statistical difference between the segments for each mode. This determined that the association between mode and segment for two years of smart card data was statistically significant ($\chi 2$ (6) =149.787, $p<.000$). A Cramer's V test identified that this relationship had a value of 0.077, indicating a very weak association between variables.

Next, the smart card data was re-analysed using the verification buffer approach detailed above. This process helped to verify the appropriate segment for each smart card ID. The results of this verification are provided in Table 5.2.

Table 5.2 - User Breakdown by Mode into Customer Fluctuation Segment over 2 years, WITH verification buffer

	Train		Bus		Tram	
	No. of users1	% of users	No. of users1	% of users	No. of users1	% of users
New	143	3.6%	368	9.9%	179	5.3%
Lost	1,982	50.0%	2,149	57.9%	1,895	56.5%
Retained	152	3.8%	71	1.9%	88	2.6%
Returning	1,688	42.6%	1,122	30.2%	1,190	35.5%
Total	4,000	100%	3,710	100%	3,352	100%

Note: All figures are based on a two-year period between October 2016 and September 2018.
These figures have been adjusted with a 3-month verification buffer at the start and end of the assessment period to confirm new and lost user rates resulting in an 18-month assessment window.
All modes had a randomly selected sample of approximately 4000 smart cards.
1Users have been defined as an individual smart card ID

Performing the customer fluctuation segmentation on an 18-month assessment period with verification at the beginning and end of the period (Table 5.2) found:

- The lost rider segment dominates for all modes at over half of the card population (50% train, 58% bus, and 57% tram). The bus is responsible for the highest number of lost smart cards over the 18-month period and as with the complete two year analysis, the train has the lowest share of lost smart cards.
- Returning users are the second largest segment for all modes, although substantially larger share for the train (43%) followed by lesser shares for the bus and tram (30 – 35%).
- Retained users remained the smallest proportion of the population for both the bus and tram (2% bus and 3% tram). This was only marginally higher for train users at 4%.

- The bus had the highest proportion of new users at 10% of all smart cards. This was nearly double the rate for the tram (5%) and higher than the rate for the train (4%).

Overall the train appeared to lose fewer users and keep more users on a returned or retained basis than the bus and tram, which showed similar patterns in user segments at a lesser scale. The bus attracted the highest proportion of new users and had the highest rate of lost users. This could be reflective of the type of users on the bus, for example, school users with frequent semester breaks. However, it might also indicate an issue with market leakage. However, based on both the two year and one year new user results for the bus, the high share of new users in this instance is more likely to be caused by a data anomaly than a meaningful change in behaviour. This information provides insights that can allow for further targeted investigation and the ability to identify whether there are service issues that increase the rate of lost bus users.

To understand the statistical significance of these results, a Pearson Chi-Square test was completed. This test identified a statistically significant relationship between the customer fluctuation segments and each individual mode when using the verification buffer ($\chi2$ (6) =260.577, p<.000), and had a Cramer's V value of 0.109. Though this is still a weak association, this value suggests a more substantive relationship between mode and segment than the two year results.

Table 5.3 provides a direct comparison of the segmentation results for the two-year assessment and the two year assessment with verification buffer.

Table 5.3 - Comparison of Two Year and Two Year with Buffer Segmentation Results

	TRAIN		BUS		TRAM	
	Two-year results	**Buffer results**	**Two-year results**	**Buffer results**	**Two-year results**	**Buffer results**
New	2.7%	3.6%	2.7%	9.9%	2.4%	5.3%
Lost	43.1%	50.0%	54.7%	57.9%	52.9%	56.5%
Retained	2.2%	3.8%	1.2%	1.9%	1.4%	2.6%
Returning	52.0%	42.6%	41.4%	30.2%	43.2%	35.5%
Total	100%	100%	100%	100%	100%	100%

Note: All modes had a randomly selected sample of approximately 4000 smart cards.

Overall the most significant change appears to be a decrease in the proportions of returning users when the buffer is applied as they can be more appropriately identified as new, lost or retained cards. The proportion of new users across all modes is higher when using the verification buffer, indicating that over an additional 3 months many of these new users are reclassified (lost or returning) due to variations in travel patterns. It is evident that when using months as the temporal unit of analysis the change within this segment will continue to be sensitive to the start and end of the measurement period. This sensitivity may become less relevant where customer fluctuation is used as a repeated measurement over several years.

There was also an increase in the proportion of lost users when utilising a verification buffer for all modes, meaning that a proportion of users (7% increase for train, 8% increase for bus and 6% for the tram) return to travel within the 3-month buffer period. Finally, there was a small increase in the proportion of retained users for all modes when utilising the buffer period. This provides some evidence of the impact of time scale on public transport user retention and mimics the theories of customer lifetime value (Reinartz and Kumar, 2000).

This raises the key question: which of the two methods is more accurate? On balance, it is suggested that the verification buffer acts to test for changes in usage patterns at the fringe of the analysis period, which is a known area of sensitivity in the analysis. Without the buffers, the two year data does not account for changes in use either before or after the analysis period. **Overall, it is considered that the use of two-year data with a verification buffer is beneficial as it most accurately accounts for customer fluctuation.**

5.4.2 Comparison of Two Year and One Year Smart Card Results

As previously mentioned, this chapter seeks to refine the measurement approach using one year of smart card data as completed in Chapter 4. Table 5.4 shows the difference in segment size between one year of smart card data, two years of smart card data and two years of smart card data with a verification buffer.

Table 5.4 - Comparison of 1-year, 2-year and 2-year with buffer customer fluctuation segmentation results

	Train			Bus			Tram		
	Smart Card 1-year[1]	Smart Card 2-year[2]	Smart Card 2-year (with buffer)[3]	Smart Card 1-year[1]	Smart Card 2-year[2]	Smart Card 2-year (with buffer)[3]	Smart Card 1-year[1]	Smart Card 2-year[2]	Smart Card 2-year (with buffer)[3]
New	21.9%	2.7%	3.6%	23.1%	2.7%	3.6%	20.4%	2.4%	5.3%
Lost	46.4%	43.1%	50.0%	51.8%	54.7%	62.4%	53.5%	52.9%	56.5%
Retained	6.7%	2.2%	3.8%	5.4%	1.2%	2.1%	4.1%	1.4%	2.6%
Returning	25.0%	52.0%	42.6%	19.8%	41.4%	32.0%	21.9%	43.2%	35.5%
Total	**100.0%**	**100.0%**	**100.0%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Note: [1] One year data measures a large sample between October 2016 – September 2017
[2,3] Two year and two-year with verification buffer data measures a smaller sample (4000 per mode) between October 2016 and September 2018

This analysis highlights a significant difference between the one year and two year smart card studies. Notably the biggest differences are in the rates of new and retained users (lower in the two year data) and the rates of returning users (higher in the two year data). The size of the lost user share generally remained consistent across all three measurement approaches. Although some of the reduced shares of new and retained users can be considered as a function of time, it is still considered that the two year smart card data with verification buffer provided the most accurate measure of customer fluctuation. As lost users were not affected or increased when using the two year method, this segment is investigated further (see section 5.4.6).

A final observation from these tests is that it is clear segment shares are very sensitive to the assumptions used in defining the analysis period. However, because the two-year data with a verification buffer was identified as the best time span for measuring customer fluctuation, this data is used for most of the analysis within this chapter. The exception is the following section which is exploring customer fluctuation by card type where the complete two year data is used to utilise the larger sample size for greater depth.

5.4.3 Customer Fluctuation by Card Type

A limitation of smart card data is that it provides minimal demographic data to allow for the analysis of patterns and differences. Although this is an important component in ensuring smart card users anonymity (for privacy reasons), it hinders the ability to review demographic patterns in travel behaviour. To address this, the two-year smart card data included some additional information, including the card type. These card types, though general, allow for a high-level understanding of how different types of users travel.

There are many different card types available within the myki system. For ease of analysis, these were grouped into the following categories: full fare cards, child cards, student cards, senior cards, and all other concession cards. Broadly, we can assume that full fare cards are used by those aged 18 – 65, child cards are for users 4 – 16, student cards are generally for those 12 – 24, and senior cards are for users that are 65 years and older. There are no age brackets that can be assumed for all other card types.

This division by card type found the following breakdowns for each mode, as shown in Table 5.5:

Table 5.5 - Breakdown of Card Type (simplified) by Mode

	Train		Bus		Tram	
Full	2,532	63.9%	1,954	55.7%	2,235	67.7%
Child	448	11.3%	484	13.8%	262	7.9%
Student	134	3.4%	165	4.7%	115	3.5%
Seniors	268	6.8%	311	8.9%	272	8.2%
Others	583	14.7%	595	17.0%	415	12.6%

This shows that:

- The tram has the highest proportion of full fare users (68%), and the bus has the lowest (approx. 56%) with the train roughly in the middle of this.

- The bus had the highest proportion of users travelling on a child fare (14%), closely followed by the train (11%). The tram had a much lower proportion of child fare tickets (8%).
- The proportion of student users was roughly consistent across all modes at between 3 – 5%.
- The share of senior users was generally consistent across all modes, though highest for bus and tram (8 - 9%), followed by train users (7%).
- All modes had roughly similar proportions of other concession card users (12 – 17%), with the rate being the highest for bus users and lowest for tram users.

These proportions should be considered when exploring the following analysis of customer fluctuation segment share by card type for each mode.

Table 5.6 provides the results of customer fluctuation segments by card type over two years using months as the unit of temporal analysis and excluding the verification buffer.

Table 5.6 - Customer Fluctuation Segments using 2-year data by the proportion of USERS per Card Type by mode

Card Type	Mode	New	Lost	Retained	Returning
Total	Train	2.7%	43.1%	2.2%	52.0%
	Bus	2.7%	54.7%	1.2%	41.4%
	Tram	2.4%	52.9%	1.4%	43.2%
Full Fare	Train	2.2%	43.9%	2.1%	51.8%
	Bus	1.8%	58.0%	0.8%	39.5%
	Tram	2.1%	53.4%	1.2%	43.3%
Child	Train	2.4%	39.0%	0.7%	57.9%
	Bus	3.6%	54.2%	1.2%	41.0%
	Tram	0.8%	61.1%	0.0%	38.2%
Student	Train	7.4%	55.1%	4.4%	33.1%
	Bus	2.3%	53.4%	2.3%	42.0%
	Tram	1.7%	54.8%	0.9%	42.6%
Senior	Train	4.0%	30.5%	3.3%	62.1%
	Bus	4.4%	39.9%	3.4%	52.3%
	Tram	5.5%	34.9%	4.8%	54.8%
Other	Train	3.1%	46.1%	2.7%	48.1%
	Bus	4.2%	52.0%	1.4%	42.3%
	Tram	3.4%	56.1%	1.9%	38.6%
Note: This assessment uses the two-year smart card data, taken from between October 2016 and September 2018					
Legend:		5% higher than for the total sample	5% lower than for the total sample		notable exceptions

A Pearson Chi-Square identified a statistically significant relationship between the customer fluctuation segments and card type, $\chi 2$ (12) =161.787, p<.000. The key insights from this analysis include:

- Across the board, student tickets had a higher proportion of lost ridership when compared to the overall two year data. This might be due to semester and non-semester time, as well as a proportion of the population that will age out of a student ticket over the period of analysis.

- Conversely, senior concession tickets had the lowest proportions of lost users at between 30 – 40% for all modes. Senior bus users had the highest rate of lost cards when compared to other modes, which is consistent with the bus having the highest share of senior cards. Senior ticket types also had the highest rates of returning users across all modes.
- Train users had the lowest proportion of lost users travelling on a child ticket (predominant in those aged 4 – 16) at 39% while tram users had the highest proportion of lost child tickets (61%), with the bus share lying in between these two modes.
- The tram had the lowest proportions of student and child card types, and very few of these travelled as retained users.
- Full fare bus users had the highest rate of lost ridership compared to both full fare cards in other modes and all other card types that travelled by bus. Full fare bus users also had the lowest rate of retained use compared to both the full data set and other modes. This might indicate that the primary audience for bus use is not those using it for commuting travel.

Further, Table 5.7 provides the results of the share of **trips** for each customer fluctuation segment by card type over two years. This also used months as the period for analysis.

Table 5.7 - Customer Fluctuation Segments by the Proportion of TRIPS per Card Type by Mode

Card Type	Mode	New	Lost	Retained	Returning
Total	Train	12.4%	33.4%	19.6%	34.5%
	Bus	14.0%	37.6%	11.4%	37.0%
	Tram	12.7%	33.9%	15.6%	37.8%
Full Fare	Train	12.3%	31.6%	20.2%	35.9%
	Bus	9.6%	39.8%	9.6%	41.1%
	Tram	13.9%	31.0%	12.4%	42.7%
Child	Train	13.4%	36.0%	8.6%	42.0%
	Bus	14.1%	45.4%	9.1%	31.5%
	Tram	3.8%	42.2%	0.0%	54.0%
Student	Train	13.1%	45.0%	18.1%	23.7%
	Bus	12.5%	47.6%	10.2%	29.6%
	Tram	2.0%	49.7%	2.3%	46.1%
Senior	Train	11.2%	30.3%	23.0%	35.4%
	Bus	21.7%	27.8%	25.9%	24.6%
	Tram	14.1%	16.9%	41.9%	27.1%
Other	Train	12.7%	34.8%	20.4%	32.0%
	Bus	11.9%	49.7%	14.7%	23.8%
	Tram	3.4%	56.1%	1.9%	38.6%
Note: This assessment uses the two-year smart card data, taken from between October 2016 and September 2018					
Legend:	(green) 5% higher than for the total sample	(orange) 5% lower than for the total sample	(blue) notable exceptions		

Overall, very different patterns are evident with trip based results compared to the card based (notionally a person based) analysis. Notably this analysis illustrated that:

- All customer fluctuation segments for bus users with a senior smart card were responsible for a roughly equal share of trips. This was the only card type and mode for which this was the case.
- Student, child, and other concession types all had a higher share of lost trips for the bus and tram (train for student only), compared to the total sample. The tram had a lower share of trips from lost senior card types.
- Overall retained tram users had a low share of trips for the child, student, and other concession card types. However, retained senior tram users had a significantly higher share of trips than for the total sample (+26% of trips)
- Child card types had a higher share of returning user trips for both the train and tram, though a lower share of trips for bus users. This is interesting to note given the knowledge that the bus sample had the highest share of child card types.

These findings provide some high-level insights into the demographic patterns of customer fluctuation segments. It is apparent that customer fluctuation segments are impacted by card type. **As understanding demographic patterns is important for marketing, further research is recommended on understanding how demographic patterns impact customer fluctuation**. This information will be derived through primary data collection via an online survey discussed in Chapter 6.

5.4.4 Total Trip Volume Analysis

Most of the measurement results so far in this thesis have been based on card users which might, at least notionally, be equated to people. The following analysis concerns trips made as to the unit of analysis rather than cards (people). As identified in the first part of Chapter 4 and by the analysis above (Section 5.4.3), the trip volume is an indicator of the total scale of usage in each market segment. A comparison of the proportion of smart cards to the proportion of total trips was completed for the two-year analysis (including the verification buffer) in Table 5.8.

Table 5.8 – Comparison of Share of Population to Total Trips for Each Customer Fluctuation Segment Over Two Years

	TRAIN		BUS		TRAM	
	% of smart cards[1]	**% of trips[2]**	**% of smart cards[1]**	**% of trips[2]**	**% of smart cards[1]**	**% of trips[2]**
New	3.6%	12.4%	2.7%	14.0%	2.4%	12.7%
Lost	50.0%	33.4%	54.7%	37.6%	52.9%	33.9%
Retained	3.8%	19.6%	1.2%	11.4%	1.4%	15.6%
Returning	42.6%	34.5%	41.4%	37.0%	43.2%	37.8%
Total	**100.0%**	**100%**	**100%**	**100%**	**100%**	**100%**

Note: All figures are based on a two-year period between October 2016 and September 2018. These figures have not been adjusted using 6-month buffer verification approach system to allow for comparison with one-year results.
[1] It is assumed for this assessment that one smart card is equivalent to one user
[2] Trips have been determined as any trip (touch on and confirmed touch off or time out) as such, continuation trips are excluded from this count.

Table 5.8 provides the following key insights:

- Lost users were the largest group of cards for all modes (50 – 55%) and were closely followed by returning users (40 – 43%). Both segments were responsible for a similar proportion of trips between 33 – 40%. All modes had a roughly equal proportion of trips for lost and returning users. While the lost segment represents the largest group of cards (50-53%), it only represents about 33-38% of trips.
- Unlike the findings of the one-year analysis, retained users were not responsible for the largest volume of trips. The retained segments for both the train and tram were responsible for the third-highest proportion of trips, at 20% and 16%, respectively. For bus users, the retained segment accounted for the lowest volume of trips at 11% of all trips taken.
- New users were a small share of both cards and trips. This is consistent with the findings of the one year study.

The findings of the total travel volume analysis identified some key areas of difference when using a two-year measurement period (including a verification buffer) compared to the initial one-year analysis. This is shown in Table 5.9.

Table 5.9 - Comparison of One-Year and Two-Year Smart Card Trip Volume Analysis

Train				
	One Year		Two Year	
	% of smart cards[1]	% of trips[2]	% of smart cards[1]	% of trips[2]
New	21.9%	23.2%	3.6%	12.4%
Lost	46.4%	22.0%	50.0%	33.4%
Retained	6.7%	39.1%	3.8%	19.6%
Returning	25.0%	15.7%	42.6%	34.5%
Bus				
	One Year		Two Year	
	% of smart cards[1]	% of trips[2]	% of smart cards[1]	% of trips[2]
New	23.1%	26.8%	2.7%	14.0%
Lost	51.8%	26.5%	54.7%	37.6%
Retained	5.4%	32.1%	1.2%	11.4%
Returning	19.8%	14.6%	41.4%	37.0%
Tram				
	One Year		Two Year	
	% of smart cards[1]	% of trips[2]	% of smart cards[1]	% of trips[2]
New	20.4%	23.8%	2.4%	12.7%
Lost	53.5%	27.0%	52.9%	33.9%
Retained	4.1%	30.0%	1.4%	15.6%
Returning	21.9%	19.2%	43.2%	37.8%

Note: 1 It is assumed for this assessment that one smart card is equivalent to one user
2. Trips have been determined as any trip (touch on and confirmed touch off or time out) as such, continuation trips are excluded from this count.

This highlights the following key changes:

- The share of trips in the new segment almost halves in the two-year analysis compared to the one-year analysis. The share of trips in the retained user segment also declines substantially.
- The share of trips in the lost and returning user segment increase substantially in the two-year analysis.

There is also an interesting observation to make about the share of trips in the retained segment; this generally halves in the two-year analysis compared to the one year analysis. This could be an indicator of a natural decline in the size of the retained market over time, the result of the measurement approach or a combination of the two.

5.4.5 Exploration of New Riders and Card Initialisation Date

The extended data set includes a 'card initialization date' which might assist in better understanding of the impact of card renewal on segment size estimates. This section investigates the potential of this information for giving insight into the proportion of new users and how this relates to card initiation date. Melbourne's smart card system requires that all smart cards expire after a period of four years. Unfortunately, these new cards do not continue with the same smart card ID. This raises the question of whether a proportion of new users are retained users that have been required to purchase a new smart card with a new ID number or are retained to other preferred modes of public transport.

New riders are a small proportion of the total riders in the sample for each mode. This can be partially attributed to the requirements for new riders to start months into the measurement period and ride consistently once appearing rather than becoming lost or returning riders. Regardless, new users were reviewed in accordance with the associated card initialisation date. The purpose of this assessment was to identify whether the first trip aligns with the card initialization date. For this assessment, this was taken as anyone travelling within two months of the card initialisation date. Though this may be a generous time to allow for card activation, this has been considered necessary to account for all types of public transport users and different frequencies of travel. The findings of this assessment are provided in Table 5.10.

Table 5.10 – Occurrence of New Riders with New Cards using the Card Initialisation Date

	TRAIN			BUS			TRAM		
	New Card IDs[1]	**Card Initialisation Date Aligns with first trip[2]**	**% of new users using new cards**	**New Card IDs**	**Card Initialisation Date Aligns with first trip**	**% of new users using new cards**	**New Card IDs**	**Card Initialisation Date Aligns with first trip**	**% of new users using new cards**
Sample of New Users per mode	107	62	57.9%	368	65	17.7%	179	37	20.7%

Note: [1]This assessment only utilised those segmented as new users as per the two-year data analysis with the verification buffer
[2] Users were identified as using a new card if their first trip was within two months of the card initialisation date provided.

This indicates that:

- Train users that had an initialisation date within two months of their first trip accounted for over half (58%) of the new user sample.
- This was dramatically higher than for bus and tram users, where new users on a new card accounted for approximately a fifth of the new user sample (18% for bus and 21% for tram).

This might provide evidence that though the train had the lowest share of new users, those that are new are more likely to be travelling on a new myki card. This indicates that new train users are more likely to be new to public transport travel when compared with the bus and tram. Unfortunately, there are several limitations to this analysis. It is not possible to discern if card initialization is because a new user is purchasing a new smart card for the first time or replacing an expired card. Overall this data provides no comprehensive means of discerning the validity of new riders, though it is valuable for providing some initial insights into the complexities of measuring market entries. If further information became available, it would be beneficial to develop an understanding of the rate of new users compared to new cards based on smart card data. In the future, authorities might record card numbers between old and new cards when cards are replaced.

What is not apparent from this assessment is what proportion of these new cards are due to issues such as lost cards or card expiry, rather than the addition of users that are new to public transport. This remains an issue limiting the measurement of customer fluctuation using smart card data. **Other methods of data collection should be investigated to help refine the measurement of customer fluctuation and ensure the appropriate estimation of new users.**

5.4.6 Exploration of Lost Riders and Card Expiry Date

Lost users are of particular interest to this study due to the high proportion found within each public transport market across all segmentation methods used. The card expiry date was included in the extended data set and used to help understand the potential rate of users that are segmented as 'lost' due to card expiry rather than a change in travel behaviour. As previously mentioned, other studies of smart cards such as Chu (2015) excluded cards that expired during the measurement period from the data being used. In this study, expired smart cards have not been excluded from the assessment, as it is important to identify the rate at which they may impact on the measurement of lost users.

The following assessment identifies the number of lost users by card type and how many of these lost smart card users stopped travel within two months of the smart card expiry date. Two months was chosen to allow time for those switching to a new smart card just before expiry. The findings of this assessment are provided in Table 5.11.

Table 5.11 - Assessment of Lost User Smart Cards that Expired During the Measurement Period using two-year data WITH buffer

	TRAIN			BUS			TRAM		
	Lost Card IDs[1]	Card expires[2]	% of lost cards that expired[3]	Lost Card IDs[1]	Card expires[2]	% of lost cards that expired[3]	Lost Card IDs[1]	Card expires[2]	% of lost cards that expired[3]
Number of Lost Users per Mode	1982	123	6.2%	2149	352	16.4%	1895	40	2.1%

Note: [1] Lost card IDs are all those segmented as lost using the two-year data with verification buffer
[2] The card expiry date is calculated as four years from the date of card initialisation (provided)
[3] Cards were identified as being lost due to expiry if they stopped travelling in the month before or month of card expiry

The key findings of this analysis are as follows:

- For bus markets (which had 55% lost riders, as shown in Table 5.8), 16% of lost cards had expired. This might explain why the proportion of lost riders was higher than for train and tram markets.
- Lost train users (50% of the train sample) had 6% lost cards due to expiry, and lost tram users (57% of the tram sample) had 2% of these losses due to card expiry.

The implication of this analysis is that the share of lost users might be adjusted to reflect card expiry. The degree of adjustment is however unclear. If all expired cards were reissued, then all expired lost users might be re-segmented as retained or returning users. However, we cannot tell what percentage of users might remain lost to a transport mode.

5.5 Discussion

The main research objective for this chapter was to explore a revised approach measuring customer fluctuation using smart card data. In Chapter 4, the first application of the customer fluctuation concept to public transport data using one year of smart card data was completed. This chapter provided an opportunity to explore a more detailed data source and varied approaches to customer fluctuation measurement to build on the initial results achieved.

5.5.1 The Measurement Approach

Overall, this analysis investigated several potential refinements for the measurement of customer fluctuation using smart card data. This included the use of an extended measurement period and verification buffer, an investigation of card type, an exploration of trip volume by segment and a review of card initialisation (and associated expiry dates) on customer fluctuation results.

The purpose of investigating an extended measurement period and the use of a verification buffer is to reduce the impacts of seasonality and timing on customer fluctuation results. Both the two year (without verification) and two years with verification buffer approaches produced similar results. For all modes, the majority share of cards are classified as lost or returning, with only a small number of cards classified as new or retained. An increase in lost and returning users is understandable when compared to the one-year measurement approach, as it is increasingly likely that users will be absent for a partial or extended period as the length of the measurement period increases.

The key difference between the two-year and verification buffer approaches is a decrease in the proportion of the returning segment when using the verification buffer and a slight increase in the proportion of new, retained and lost segments. This is beneficial as we have a greater proportion of cards in the more distinct segments of new, lost, and retained. There is also, due to the manual verification process, a greater certainty that users that appear at the end of the measurement period (and to a lesser extent the beginning) have been segmented to best reflect their ongoing travel patterns. The reduction in the share of returning user segments might also provide evidence that the verification buffer approach allows sufficient time to adjust for the potential impact of seasonality. **Based on these findings, the verification buffer is identified as a preferred approach for future studies.**

To address another limitation of smart card data, the use of card types to derive high-level demographic patterns in customer fluctuation behaviour was investigated. Across all modes, student tickets had a higher proportion of lost ridership, which might be due to the start-stop nature of travel associated with school semesters or merely a reflection of the transition from student to full fare cards. Conversely, senior ticket types had the lowest proportion of lost users (30 – 40%) for all modes, and the highest proportions of returning users (50 – 60%) which might provide evidence that senior uses are likely to travel intermittently for leisure or appointments, as they no longer need to travel to commute.

The total travel volume (trips) was also investigated and compared between the one year and two-year smart card data. This found that in the two-year measurement approach the proportion of trips for new users and retained users both decrease while the proportion of trips for returning users increase. These findings indicate that the value of customers for public transport markets (frequency of ridership) may decline rather than increase with time. This argument is interesting in that it contradicts the theories of customer retention resulting in continuous increases in patronage, though does reflect the idea of customer lifetime value, where even loyal customers are seen to reduce or cease patronage at a certain point in time (Reichheld, 1996, Reinartz and Kumar, 2003).

When comparing the proportions of card types and how they are segmented between modes some further insights are gleaned. This includes the train being attractive to those travelling on a student ticket, with a high proportion of retained and new student tickets. This may be influenced by a key cluster of higher education institutions in or near the city or located in outer suburbs. In comparison, the bus had the highest rate of lost ridership from full fare card types and the lowest retained rate for these users. However, due to the generalisations required by the card type segments, these insights are preliminary observations at best.

Finally, this chapter investigated the card initialisation date to understand new and lost segments and how expired or new smart cards might impact segment share. For new users, both train and bus had over half of the new users identified using a card that had an initialization date that aligned with the first date of card use, and for tram users it was slightly less than half. However, it is then difficult to discern which of these new cards are legitimately new cards; some may be a renewal of old cards.

The card expiration date was also used to determine whether lost users were overly identified due to the influence of expired smart cards. Only a small share of lost user cards aligned with a card expiry date (except for the bus market at 15% of cards). This implies that a small proportion of lost users may simply be changing over to a new smart card.

5.5.2 Opportunities for Refinement

The use of the extended measurement period was considered to provide an improved measurement that accounts for the impact of seasonality. However, there is much room to explore seasonality in greater detail.

The use of card type to provide demographic information was insufficient to provide detailed demographic data. This limitation is difficult to address due to concerns around privacy with smart card users. This is an existing issue with both smart card data (Dempsey, 2015) and churn measurement more generally (Holtrop et al., 2016). Without additional associated data, for example, a method to connect smart card IDs with individual users, smart card data alone may not be the best data source for measuring customer fluctuation.

Further, the use of smart card data provides no insights as to *why* customer fluctuation groups might be exhibiting certain types of behaviour. To extend the findings of this exploration using smart card data, there is a clear need for analysis using primary research methods. This is the focus of the next chapter. Despite the identified limitations, the use of smart card data to measure customer fluctuation does offer a relatively efficient approach to gain quick insights into internal variability of public transport ridership.

5.6 Conclusion and Next Steps

This chapter investigated refined approaches to the measurement of customer fluctuation, based on the findings of Chapter 4. Though improved approaches were found, there are still evident limitations with the ability of smart card data to capture meaningful insights in relation to customer fluctuation. In simple terms, smart cards represent a limited means to understand human behaviour. The primary limitation is the requirement for anonymous data, which means researchers cannot trace, contact, or interact with individuals based on their smart card ID. It is not possible to question card users around their travel patterns, which might allow for a better understanding and refinement of the customer fluctuation measurement approach. As such, methods for primary data collection are suggested to help refine the measurement of customer fluctuation and ensure the appropriate estimation of new users. This is the focus of the next chapter of this thesis.

Chapter 6: Measuring Customer Fluctuation Using a Cross-Sectional Survey

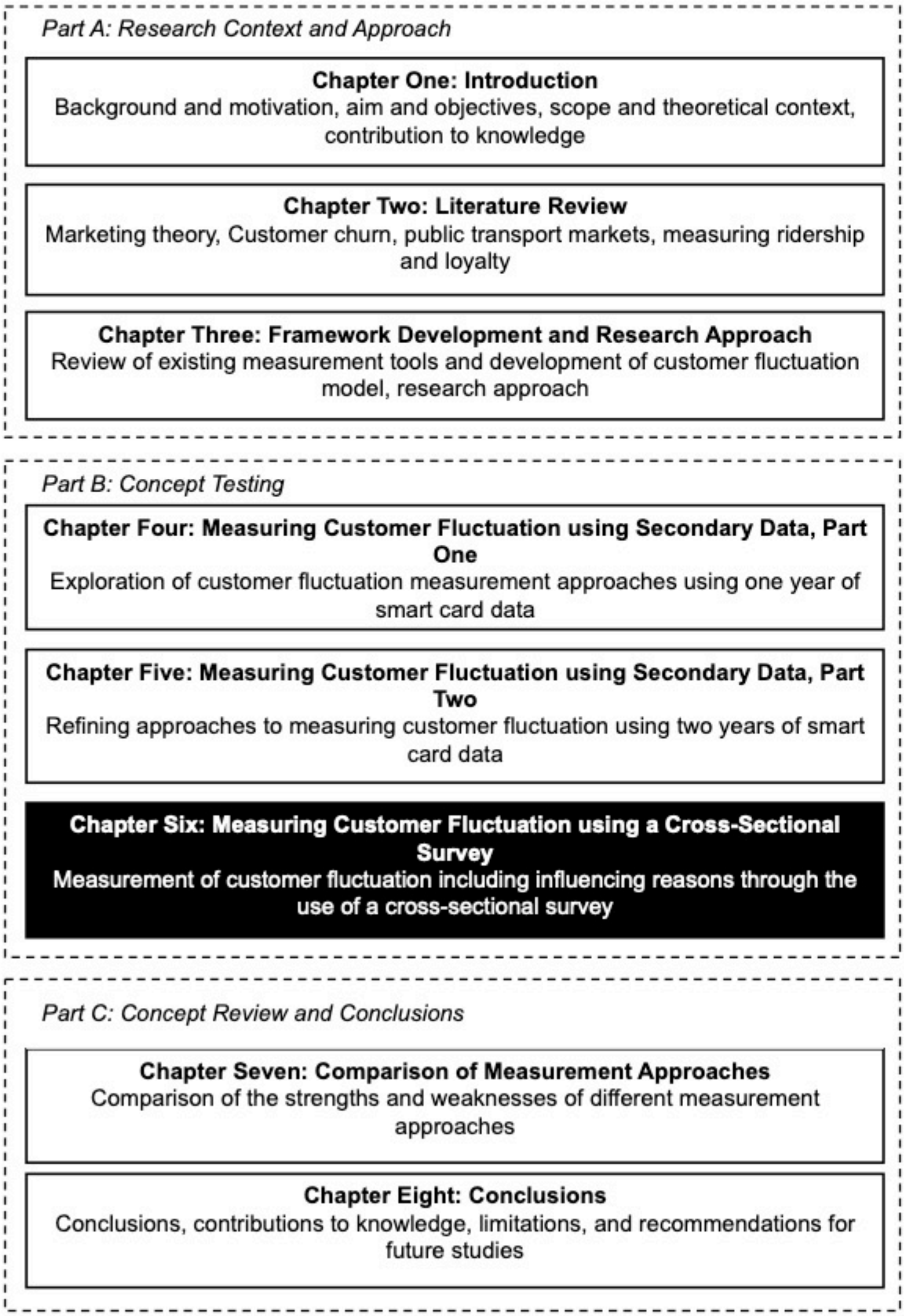

Figure 6.1 - Position of Chapter 6 in thesis structure

6.1 Introduction

This chapter presents the findings from the cross-sectional survey that collected primary data to measure customer fluctuation within Melbourne public transport markets. This will be referred to as the 'customer fluctuation survey' or the 'survey' for the remainder of this chapter. The survey was conducted between November 2018 and February 2019 and is the only primary data collected as part of the mixed-method approach discussed in Chapter 3: Research Approach. The purpose of using the customer fluctuation survey is to explore an alternative approach to the measurement of customer fluctuation and to identify the reasons influencing customer fluctuation behaviour among participants. The customer fluctuation survey provides insights into why customer fluctuation is occurring that cannot be obtained using smart card data.

This chapter begins by stating the aims of the analysis and briefly reiterating the research approach. This is followed by the measurement of customer fluctuation segments through the survey. Then a review of the socio-demographics within customer fluctuation segments is undertaken. The core section of this chapter is the exploration of the reasons behind customer fluctuation behaviour. The final section provides a discussion of the results and outlines the key conclusions from this analysis.

6.2 Research Aims and Objectives

The primary data analysis undertaken in this chapter continues the focus on addressing Research Aim I. This aim is re-iterated below for reference:

> I. To develop, measure and apply a new concept for market change analysis based on market change segments.

The applicable research objectives being investigated within this chapter are:

> *RO3. To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne*
>
> *RO4. To explore behavioural factors influencing market change segments using survey data for metro Melbourne*

The research approach for answering these questions is summarised in the next section. Full details of the research approach are provided in Chapter 3.

6.3 Research Approach

This analysis aims to measure the size of customer fluctuation segments within Melbourne public transport markets and understand the reasons behind behaviours of each segment. The analysis also seeks to investigate whether there are any links between socio-demographic patterns and customer fluctuation segments.

6.3.1 Survey Method

The customer fluctuation survey is an online survey that samples both public transport users and former public transport users within the Melbourne Metropolitan Area. The survey targets users from all modes (train, bus and tram) as well as previous users of public transport. A retrospective approach has been utilised to capture longitudinal behaviour in the survey. This has been identified to have sufficient accuracy for this type of research (Beige and Axhausen, 2008); regardless, questions were simplified to reduce any potential issues with memory recall.

The survey starts with a series of demographic screening questions to fit a sample frame which is designed to be generally representative of Melbourne pubic transport users. This is followed by asking participants to identify their main travel mode, which became the focus of the rest of the survey for that participant. Participants were not required to answer questions about secondary modes they use for travel.

A series of questions were then used to measure the participant's customer fluctuation segment. This used a two-prong approach: participants were first asked to identify whether they had travelled a minimum of once for each month of the previous year (hereafter called their 'verified' segment); they were then asked to categorise their travel based on a series of descriptive statements reflecting each customer fluctuation segment (hereafter called their 'self-selected' segment). As introduced in Chapter 3, the four statements that respondents chose from were:

1. *I started using the train/tram/bus a few months into the year and used it most months once I started* **(new user)**
2. *I used the train/tram/bus for at least one month this year and then didn't use the train/tram/bus again* **(lost user)**
3. *I used the train/tram/bus every month, or most months* **(retained user)**
4. *I used the train/tram/bus sometimes* **(returning user)**

This method makes it possible to compare the accuracy of the self-selected segmentation approach against verified travel behaviour patterns.

Based on their self-selected segmentation, participants identified the top three reasons that influenced their choice of self-selected segment, to identify the reasons behind customer fluctuation behaviour. Participants were asked to provide any further comments clarifying their reasons for customer fluctuation behaviours via an open response question.

A summary of the survey structure, which is presented in more detail in Chapter 3, is provided in Figure 6.2; the questionnaire is also provided in **Appendix A** of this thesis**.**

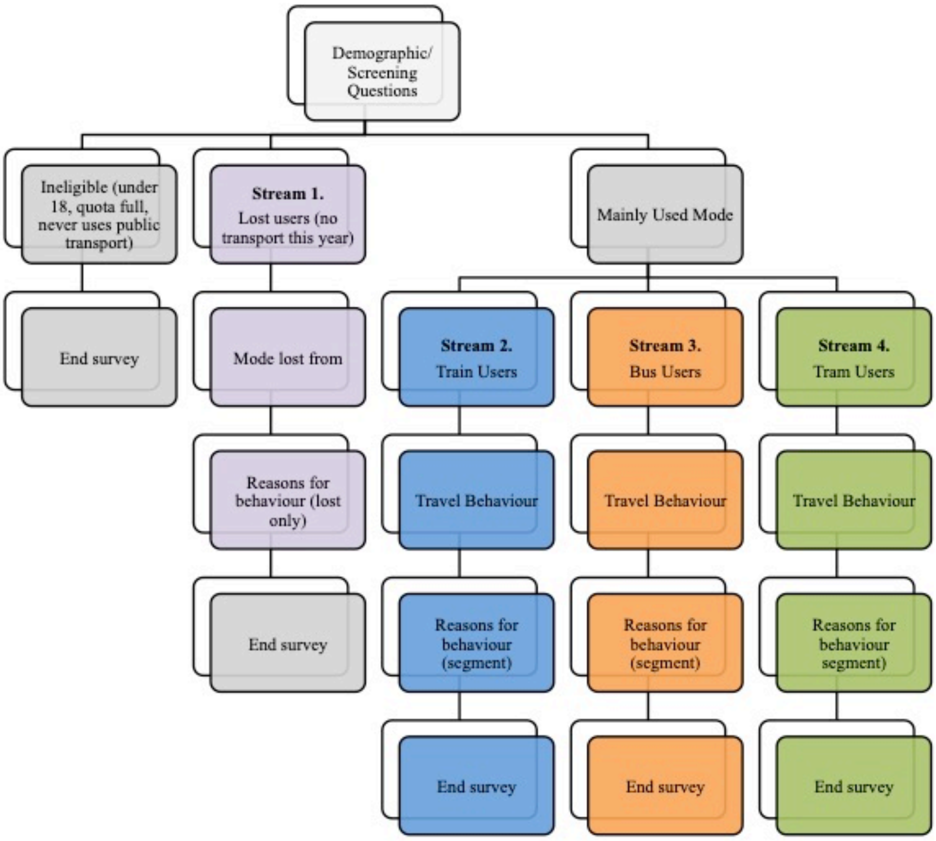

Figure 6.2 - Questionnaire Structure Highlighting the Four Streams

Data and Sampling

A sample frame was developed to be broadly representative of the population of Melbourne. It was determined by age, education and income, using response quotas. The target for sampling was 400 users per mode. This sample size would achieve a 95% confidence interval (CI) +/- 5% given the population of Melbourne, at the time of the survey.

The survey was managed and principally collected by a commercial market research company (IPSOS). Where there were low response rates, additional data collection was undertaken by the research team. The survey platform 'Qualtrics' was used to collect and store responses. Participants responding through the market research survey link were provided with a small payment for their time in completing the survey, managed independently from the research team.

The survey ran from late November to mid-December, 2018 and initially produced a sample of 430 train users, 338 tram users and 190 bus users. This was below the quota for tram and bus. Additional sampling was conducted for bus and tram users in March of 2019 by 'pushing' the survey link on social media sites Facebook, LinkedIn, Twitter and a post within the Monash Staff portal. This additional sampling provided a further 61 bus users and 45 tram users. Despite the additional sampling, the bus and tram quota was not achieved, this was a limitation of the survey design requiring users to select a 'mainly used mode', which may have deterred multi-modal users from responding. Further, additional sampling had to been concluded due to the time and funding constraints of this research. As such, Confidence Intervals were modified downwards to reflect the sample size for the bus and tram, as shown in Table 6.1.

Table 6.1 - Confidence Interval for Each Mode

Mode	Sample	Confidence Interval	Margin of Error
Train	430	95%	5
Tram	346	90%	10
Bus	251	85%	15

The demographics for each sample, as they compare to the demographics of greater Melbourne are provided in Table 6.2.

Table 6.2 - Demographic Breakdown of Users Sampled

	Melbourne* (%)	Train (%)	Bus (%)	Tram (%)
Gender				
Male	49	33.5	41.8	47.4
Female	51	66.3	57.4	51.4
Other	-	0.2	0.8	1.2
Age				
18 - 19	1.5	1.4	5.2	2.6
20 - 29	15.5	18.1	31.5	33.8
30 - 39	15.5	29.5	25.5	26.0
40 - 49	13.9	24.2	15.1	13.9
50 - 59	11.9	17.4	16.3	15.6
60 - 69	9.3	7.4	5.2	5.5
70+	9.7	1.9	1.2	2.6
Employment				
Employed full-time	58	53.3	38.6	48.0
Employed part-time	30.6	21.4	21.5	19.4
Unemployed looking for work	6.8	3.7	6	5.5
Student	-	3.5	10.8	9.8
Retired	-	5.6	4.8	8.4
Home duties	-	9.5	14.3	5.5
Other	4.6	3.0	4	3.2
Household Structure				
Family with children under 18	50.3	38.1	35.1	25.4
Family with adult children	25.3	16.7	16.7	10.4
Couple with no children	23.1	19.1	17.9	28.6
Single	23.2	19.3	20.3	24.0
Group	5.0	3.7	7.2	9.8
Other	1.5	3	2.8	1.7
Income				
Negative/nil income	-	3.3	5.6	3.8
$1 - $299 ($1 - $15,548 p.a)	3.4	5.6	9.6	7.8
$300-$499 ($15,600 - $25,948 p.a)	8.6	37	34	1.7
$500 - $799 ($26,000 - $41,548 p.a)	11.0	13	15.9	12.4
$800 - $1,249 ($41,600 - $64,948 p.a)	15.5	21.6	11.6	14.7
$1,250 - $1, 749 ($65,000 - $90,948 p.a)	14.8	17.9	15.9	17.9
$1,750 - $2,999 ($91,000 - $155,948 p.a)	26.4	16.3	13.5	15.6
$3000 or more ($156,000 or more p.a)	17.8	5.3	4.4	4.9
Prefer not to say	2.5	8.4	10	11.6
Education				
Post Graduate Degree	37	21.9	19.9	15.9
Bachelor Degree		41.6	31.1	43.4
Graduate diploma or certificate	12.9	17.7	22.3	18.2
Year 10 or above	29.6	18.1	25.5	20.8
No educational attainment	1.2	0.7	1.2	1.7

*Melbourne Data from the 2016 Census of Population and Housing Community Profile for the Greater Melbourne Area (Australian Bureau of Statistics, 2018a)

Based on the above, we can see that there are a few variations worth noting within the sample. Both the train and bus samples over-represent women (66% of train users and 57% of bus users). Those aged 30 – 49 are over-represented for the train and those aged 20 – 39 are over-represented for the bus and tram. All modes slightly under-represented the population aged 60+, which is potentially a limitation caused by undertaken the survey exclusively through an online portal. The sample for all modes also includes a larger proportion of users with a bachelor’s or post graduate degree when compared to the Melbourne average which might cause some bias in results. Regardless the direction of the bias cannot be confirmed or applied to all users with a higher education. There may be some increase in pro-public transport attitudes, but Davidov (2007) argues that those with a higher education might have a lower preference for public transport use. The use of university based surveys in the public transport industry is also common (e.g. Schmitt et al., 2013). Finally, there is also an under-representation of full-time employed bus and tram users.

Some of these variations are anticipated due to the characteristics of each mode's unique market – for example, the train is a primary mode for young to mid-career professionals travelling into the city for business where as the bus market has a significant market for school travel and retirees. While the sample is far from perfect, it is generally considered appropriate for exploratory testing in the research.

Analysis Methods

The survey analysis focused on a systematic exploration of the data obtained and what new information it can provide in relation to public transport markets. The analysis presents the results of measuring customer fluctuation using both the self-selected responses from participants and verified responses from their behaviour patterns. Verified responses were obtained by applying the a-priori segmentation method developed in Chapter 3: Research Approach to the one year of monthly travel data participants provided. The results of each measurement approach are then compared. All remaining analysis uses the self-selected segmentation results and includes a review of socio-demographic patterns within customer fluctuation segment by each of the three modes and the reasons given for customer fluctuation behaviour.

Further detail about the research approach is provided in Chapter 3. The results of this analysis are now presented.

6.4 Results

6.4.1 Measurement of Customer Fluctuation

Table 6.3 - Results of Participant Self-Selection into Customer Fluctuation Segments

	Train[1]		Bus[2]		Tram[3]	
	No. of Users	% of train sample	No. of Users	% of bus sample	No. of Users	% of tram sample
New	60	14%	46	18.3%	83	21.7%
Lost	27	6.3%	27	10.8%	43	11.2%
Retained	211	49.1%	131	52.2%	177	46.2%
Returning	132	30.7%	47	18.7%	80	20.9%
Total	430	100.0%	251	100.0%	383	100.0%

The findings of the self-selected customer fluctuation segments are presented in Table 6.3 and are summarised as:

- The retained segment had the highest respondent share of all modes with bus having the most retained users (52%), followed by the train (49%) and the tram (46%).
- Returning users had the second-highest share of respondents for the train (31%) and bus (19%), although it was an approximately equal share of users with the new segment for the tram at 21% of respondents.
- New users were over a fifth of the sample for tram users (22%) followed by the bus at 18% with the train having the lowest proportion of new users at 14%.
- Lost users were the smallest segment for all modes at 11% for tram users, 11% of bus users and just 6% of train users.

In addition to self-selected segments customer fluctuation results were cross-verified using an a-priori segmentation approach applied to the one year of travel recall information collected by the customer fluctuation survey. This approach checks the quality of self-selection vs the rule based a-priori segmentation method. Table 6.4 presents the outcome of this check.

Table 6.4 - Comparison of Self-Selected and Verified Customer Fluctuation Segmentation Process

	Train		Bus		Tram	
	Self-selected	**Verified**	**Self-selected**	**Verified**	**Self-selected**	**Verified**
New	14.0%	8.1%	18.3%	10.0%	21.7%	10.9%
Lost	6.3%	16.8%	10.8%	16.5%	11.2%	7.2%
Retained	49.1%	48.4%	52.2%	50.5%	46.2%	56.8%
Returning	30.7%	26.8%	18.7%	23.0%	20.9%	25.1%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**10.9%**

Table 6.4 shows that:

- Compared to the a-priori verified results, respondents were more likely to self-select themselves as being in the new segment across all modes; +7% for the train, +8% for the bus and +11% for the tram.
- Conversely, self-selection of the lost user segment was lower than the a-priori rules implies; -11% for train, -6% for bus and -4% for tram.
- Self-selection and a-priori estimates of retained and returning segments were very similar.

The verified results show a greater similarity to the smart card data results. However, the proportion of retained users remains much higher and lost users much lower in the self-selected method. Both verified and self-selected results were analysed using a Pearson Chi Square test which found that both methods had a significant relationship between mode and customer fluctuation segment (χ^2 (6) =67.967, p<.000). Though both results produced a significant relationship between mode and customer fluctuation segment, it is considered that the verified results offer an improved accuracy and ability to compare results to those gained when using smart card data for measurement. Overall the self-selection approach appears to have weaknesses which will be explored further.

6.4.2 Socio-Demographic Patterns within Customer Fluctuation Segments

Survey data provides us with the capability to understand the socio-demographic characteristics of travellers in different customer fluctuation segments and to identify whether these segments are associated with different socio-economic patterns. Although the survey was not able to collect an entirely representative sample for any mode of transport, it can provide us with a general picture of socio-demographic patterns and how they are influenced by mode and customer fluctuation segment.

6.4.2.1 Train Users

Table 6.5 provides the breakdown of the train users by the customer fluctuation segment to identify socio-demographic patterns.

Table 6.5 – Demographics of TRAIN SAMPLE by Customer Fluctuation Segment

TRAIN **USERS**	New	Lost	Retained	Returning	Total
Gender					
Male	36.7%	18.5%	35.1%	32.6%	33.5%
Female	63.3%	81.5%	64.9%	66.7%	66.3%
Other	0.0%	0.0%	0.0%	0.8%	0.2%
Age					
18 - 19	5.0%	0.0%	1.4%	0.0%	1.4%
20 - 29	25.0%	11.1%	21.8%	10.6%	18.1%
30 - 39	31.7%	51.9%	27.5%	27.3%	29.5%
40 - 49	23.3%	25.9%	21.8%	28.0%	24.2%
50 - 59	10.0%	7.4%	19.9%	18.9%	17.4%
60 - 69	5.0%	3.7%	5.7%	12.1%	7.4%
70+	0.0%	0.0%	1.9%	3.0%	1.9%
Employment					
Employed full-time	45.0%	48.1%	63.5%	41.7%	53.3%
Employed part-time	28.3%	22.2%	16.1%	26.5%	21.4%
Unemployed looking for work	3.3%	0.0%	3.8%	4.5%	3.7%
Student	6.7%	0.0%	5.2%	0.0%	3.5%
Retired	1.7%	3.7%	4.7%	9.1%	5.6%
Home duties	13.3%	22.2%	4.7%	12.9%	9.5%
Other	1.7%	3.7%	1.9%	5.3%	3.0%
Household Structure					
Family with children under 18	35.0%	44.4%	33.2%	46.2%	38.1%
Family with adult children	23.3%	0.0%	15.6%	18.9%	16.7%
Couple with no children	11.7%	33.3%	20.4%	17.4%	19.1%
Single	25.0%	3.7%	24.6%	11.4%	19.3%
Group	3.3%	3.7%	4.3%	3.0%	3.7%
Other	1.7%	14.8%	1.9%	3.0%	3.0%
Income					
Negative/nil income	5.0%	3.7%	1.9%	4.5%	3.3%
$1 - $299 ($1 - $15,548 p.a)	6.7%	0.0%	5.7%	6.1%	5.6%
$300-$499 ($15,600 - $25,948 p.a)	11.7%	18.5%	6.6%	8.3%	8.6%
$500 - $799 ($26,000 - $41,548 p.a)	15.0%	18.5%	10.4%	15.2%	13.0%
$800 - $1,249 ($41,600 - $64,948 p.a)	23.3%	11.1%	21.3%	23.5%	21.6%
$1,250 - $1,749 ($65,000 - $90,948 p.a)	20.0%	25.9%	19.0%	13.6%	17.9%
$1,750 - $2,999 ($91,000 - $155,948 p.a)	15.0%	18.5%	19.4%	11.4%	16.3%
$3000 or more ($156,000 or more p.a)	3.3%	3.7%	7.6%	3.0%	5.3%
Prefer not to say	0.0%	0.0%	8.1%	14.4%	8.4%
Education					
Post Graduate Degree	25.0%	14.8%	24.2%	18.2%	21.9%
Bachelor Degree	46.7%	44.4%	43.1%	36.4%	41.6%
Graduate diploma or certificate	15.0%	14.8%	17.1%	20.5%	17.7%
Year 10 or above	13.3%	25.9%	14.7%	24.2%	18.1%
No educational attainment	0.0%	0.0%	0.9%	0.8%	0.7%

[1] These figures are based on the sample of 430 train user responses. Users self-selected into customer fluctuation segments.

Legend: [green] 5% higher than for the total sample; [orange] 5% lower than for the total sample.

Based on the sample of train users, we can identify the following core demographic traits:

- Compared to the total gender split of train users in the sample, Female users (+15%), were much more likely to be within the lost segment than male users (-15%). There was no other substantive difference in customer fluctuation segment based on gender.

- Compared to the total train sample, the share of new users between the ages of 20-29 was high; new users between the ages of 50 – 59 were a relatively lower share.
- For the age group, the highest share of lost users was for those between the ages of 30 – 39; this may be, at least in part, be due to the transition to parenthood of people at this age. This can also be linked to the higher share of lost users (22%) that identified they were employed in home duties.
- There was a lower share of full-time employed users within the new segment when compared to the total. However, this was offset by a higher (+10%) share of retained users that were employed full time.
- There were no lost or returning student train users. This is likely influenced by the small share of student train users within the survey sample.
- Those living alone had a higher share of new users and a lower share of lost and returning users when compared to the total population.
- Income did not significantly vary between the total sample and the new, retained or returning segments, however a higher share of lost users earned \$300-\$499, \$500 - \$799 and \$1,250 - \$1, 749 per week. There was a lower than total share of lost users earning between \$800-\$1,249 per week.

A series of Pearson Chi-Square tests were utilised to identify if there were any statistically significant relationships between the customer fluctuation segments and demographic variables. This analysis identified a statistically significant relationship between customer fluctuation segment and age (χ^2 (18) =35.019, p<.009), employment (χ^2 (18) =44.360, p<.001) and household structure (χ^2 (15) =41.521, p<.000). No statistically significant relationship was identified between segments and gender, income or educational attainment.

Overall, a significant number of train users were commuters employed full time. There was also a significant proportion of new users between 20 – 29 years, which may be indicative of a transition from education to the working for the younger workforce. Conversely, train users exhibit a higher share of lost users between the ages of 30 – 39, which may relate to the higher proportion of lost female users and lost users that live as a couple with no children or a couple with children under the age of 18. This might be an indication of the negative impact of the transition to family life on public transport use.

Finally, there was a significant decrease in the proportion of retained users between the 50 – 59 and 60 – 69 age brackets. This might indicate the impact of transitions to senior employment or retirement on the frequency of public transport use for train users. Based on these findings, we suggest that life course events may have a significant impact on the customer fluctuation segment patterns of train users.

6.4.2.2 Bus Users

Table 6.6 provides the demographic breakdown of bus users by their categorisation into customer fluctuation segments.

Table 6.6 – Demographics of BUS SAMPLE by Customer Fluctuation Segment

BUS USERS					
	New	Lost	Retained	Returning	Total
Gender					
Male	32.6%	29.6%	48.9%	38.3%	40.2%
Female	65.2%	70.4%	50.4%	61.7%	52.6%
Other	2.2%	0.0%	0.8%	0.0%	0.8%
Age					
18 - 19	4.3%	0.0%	7.6%	2.1%	5.2%
20 - 29	43.5%	40.7%	27.5%	25.5%	28.7%
30 - 39	34.8%	7.4%	27.5%	21.3%	25.5%
40 - 49	4.3%	37.0%	13.0%	19.1%	12.7%
50 - 59	8.7%	11.1%	18.3%	21.3%	15.5%
60 - 69	4.3%	3.7%	6.1%	4.3%	4.8%
70+	0.0%	0.0%	0.0%	6.4%	1.2%
Employment					
Employed full-time	52.2%	29.6%	39.7%	27.7%	38.0%
Employed part-time	19.6%	11.1%	22.9%	14.9%	20.0%
Unemployed looking for work	2.2%	3.7%	6.9%	8.5%	6.1%
Student	4.3%	3.7%	14.5%	10.6%	11.0%
Retired	2.2%	7.4%	3.1%	10.6%	4.1%
Home duties	17.4%	18.5%	9.9%	21.3%	13.5%
Other	2.2%	3.7%	3.1%	4.3%	2.9%
Household Structure					
Family with children under 18	56.5%	40.7%	25.2%	38.3%	56.5%
Family with adult children	6.5%	11.1%	19.8%	21.3%	6.5%
Couple with no children	15.2%	14.8%	19.8%	17.0%	15.2%
Single	15.2%	22.2%	22.1%	19.1%	15.2%
Group	2.2%	7.4%	9.9%	4.3%	2.2%
Other	4.3%	3.7%	3.1%	0.0%	4.3%
Income					
Negative/nil income	8.7%	11.1%	4.6%	2.1%	8.7%
$1 - $299 ($1 - $15,548 p.a)	8.7%	7.4%	10.7%	8.5%	8.7%
$300-$499 ($15,600 - $25,948 p.a)	6.5%	3.7%	14.5%	23.4%	6.5%
$500 - $799 ($26,000 - $41,548 p.a)	13.0%	7.4%	16.8%	21.3%	13.0%
$800 - $1,249 ($41,600 - $64,948 p.a)	8.7%	22.2%	9.9%	12.8%	8.7%
$1,250 - $1, 749 ($65,000 - $90,948 p.a)	19.6%	14.8%	16.8%	10.6%	19.6%
$1,750 - $2,999 ($91,000 - $155,948 p.a)	15.2%	11.1%	14.5%	8.5%	15.2%
$3000 or more ($156,000 or more p.a)	6.5%	3.7%	5.3%	0.0%	6.5%
Prefer not to say	13.0%	18.5%	6.9%	12.8%	13.0%
Education					
Post Graduate Degree	19.6%	7.4%	26.7%	8.5%	19.1%
Bachelor Degree	52.2%	25.9%	22.9%	36.2%	29.9%
Graduate diploma or certificate	19.6%	11.1%	20.6%	36.2%	21.9%
Year 10 or above	6.5%	51.9%	29.0%	19.1%	21.9%
No educational attainment	2.2%	3.7%	0.8%	0.0%	0.8%

[1] These figures are based on the sample of 251 bus user responses. Users self-selected into customer fluctuation segments.

Legend: (green) 5% higher than for the total sample; (orange) 5% lower than for the total sample

For bus users, the following key demographic differences between customer fluctuation segments were identified:

- Male bus users represented a higher than sample share of the retained user segment when compared to the total bus sample and lower rates of new and lost users. Women represented a higher share of new, lost and returning ridership when compared to the total bus sample.
- Bus users had a higher share of new users between the ages of 20 – 39 when compared to the total sample, and a lower percentage of new users from the ages of 40 – 59. However, this was coupled with a higher share of lost users for the 20 – 29 and 40 – 49 age groups. There was also a lower share of lost users between the ages of 30 – 39 when compared to the total bus sample.
- Full time and part-time workers had a lower share of lost and returning users when compared to the total sample. Full time workers also had a higher percentage of new riders.
- There were lower than average rate of new and lost student segments.
- Those living as a family with adult children, alone or in a share house situation all had a higher share of retained users when compared to the total sample.
- A higher share of both retained and returning customers for bus users within the sample earned between $300-$499 per week. There was also a higher share of returning users for those who earned $500 - $799 per week.
- Those with a post-graduate degree had a lower share of lost and returning users and a higher share of retention compared to the total sample. Similarly, those with a bachelor degree had a higher percentage of new and returning users when compared to the total sample.

Chi-Square tests were identified a statistically significant relationship between bus customer fluctuation segment and age (χ^2 (18) =41.634, p<.001) and education (χ^2 (12) =41.937, p<.000). There was no statistically significant relationship between gender, employment, household structure or income and customer fluctuation segment.

Overall, the bus market had less consistent socio-economic patterns within the sample when compared to train users. There was some evidence that the bus is more attractive to older users due to higher shares of returning users that are retired. There was also some evidence of captivity in bus users, with higher shares of retained or returning users that had a lower household income, were students or lived in single or share housing.

There is also a high proportion of lost users that have an education to a year 10 or above schooling level. A potential cause for this may be that people who work in a trade or retail may stop using the bus as other means of public transport become available – for example being able to afford a car.

When investigating the reasons behind travel decisions, we may expect to see greater evidence of captivity for bus users, for example personal factors such as ‘I don’t have a car or can’t drive’.

6.4.2.3 Tram Users

Table 6.7 provides the demographic breakdown of tram users by their categorisation into customer fluctuation segments.

Table 6.7 - Demographic Breakup of TRAM SAMPLE by Customer Fluctuation Segment

TRAM USERS[1]					
	New	**Lost**	**Retained**	**Returning**	**Total**
Gender					
Male	48.2%	40.0%	50.8%	41.8%	47.4%
Female	51.8%	60.0%	48.0%	56.0%	51.4%
Other	0.0%	0.0%	1.1%	2.2%	1.2%
Age					
18 - 19	1.8%	5.0%	2.8%	2.2%	2.6%
20 - 29	53.6%	25.0%	34.1%	23.1%	33.8%
30 - 39	25.0%	30.0%	26.8%	24.2%	26.0%
40 - 49	10.7%	15.0%	13.4%	16.5%	13.9%
50 - 59	8.9%	15.0%	14.5%	22.0%	15.6%
60 - 69	0.0%	10.0%	6.1%	6.6%	5.5%
70+	0.0%	0.0%	2.2%	5.5%	2.6%
Employment					
Employed full-time	51.8%	45.0%	49.7%	42.9%	48.0%
Employed part-time	17.9%	25.0%	19.0%	19.8%	19.4%
Unemployed looking for work	5.4%	5.0%	5.0%	6.6%	5.5%
Student	16.1%	10.0%	9.5%	6.6%	9.8%
Retired	1.8%	10.0%	8.9%	11.0%	8.4%
Home duties	5.4%	5.0%	5.0%	6.6%	5.5%
Other	1.8%	0.0%	2.8%	6.6%	3.2%
Household Structure					
Family with children under 18	28.6%	40.0%	20.1%	30.8%	25.4%
Family with adult children	8.9%	10.0%	8.9%	14.3%	10.4%
Couple with no children	21.4%	20.0%	33.5%	25.3%	28.6%
Single	28.6%	25.0%	24.0%	20.9%	24.0%
Group	12.5%	5.0%	11.2%	6.6%	9.8%
Other	0.0%	0.0%	2.2%	2.2%	1.7%
Income					
Negative/nil income	5.4%	15.0%	2.2%	3.3%	3.8%
$1 - $299 ($1 - $15,548 p.a)	7.1%	5.0%	7.3%	9.9%	7.8%
$300-$499 ($15,600 - $25,948 p.a)	7.1%	25.0%	13.4%	6.6%	1.7%
$500 - $799 ($26,000 - $41,548 p.a)	23.2%	0.0%	9.5%	14.3%	12.4%
$800 - $1,249 ($41,600 - $64,948 p.a)	17.9%	10.0%	16.2%	11.0%	14.7%
$1,250 - $1, 749 ($65,000 - $90,948 p.a)	16.1%	15.0%	19.0%	17.6%	17.9%
$1,750 - $2,999 ($91,000 - $155,948 p.a)	12.5%	20.0%	15.1%	17.6%	15.6%
$3000 or more ($156,000 or more p.a)	1.8%	5.0%	5.6%	5.5%	4.9%
Prefer not to say	8.9%	5.0%	11.7%	14.3%	11.6%
Education					
Post Graduate Degree	21.4%	10.0%	16.8%	12.1%	15.9%
Bachelor Degree	44.6%	35.0%	50.3%	30.8%	43.4%
Graduate diploma or certificate	12.5%	30.0%	12.8%	29.7%	18.2%
Year 10 or above	17.9%	15.0%	19.0%	27.5%	20.8%
No educational attainment	3.6%	10.0%	1.1%	0.0%	1.7%

[1] These figures are based on the sample of 346 tram user responses. Users self-selected into customer fluctuation segments.

Legend:	(green)	5% higher than for the total sample	(orange)	5% lower than for the total sample

For tram users, the following key socio-demographic patterns in customer fluctuation segment were identified:

- Male users had lower rates of the lost segment when compared to the total tram sample, whereas women had higher rates.
- There was a significantly higher proportion of new users aged 20 – 29 when compared to the total sample and a lower share of lost and returning segments within the same age bracket. The other significant variations by age were a lower share of new segments aged 50 – 59, and a higher share of returning users in this age bracket. There were no new segment respondents aged older than 50 – 59.
- There was a higher share of lost users that were employed part-time compared to the total sample of part-time employees. There was also a slightly higher share in the proportion of new users that were students and a lower percentage of new users that were retired.
- Those with a family with kids under 18 were more likely to be categorised as lost or returning. Comparatively, couples with no children were more likely to be retained users and were less likely to be new or lost when compared to the total sample of couples with no children.
- Those with no income were more likely to be lost users when compared to the total sample. This pattern was also observed for those who earned between \$300 - \$499 per week.
- Those with higher levels of education (bachelor or postgraduate degrees) had a lower likelihood of being a lost user when compared to the total sample, and those with a bachelor degree had a higher percentage of retained users.

Chi-Square tests were identified a statistically significant relationship between tram customer fluctuation segment and age (χ^2 (18) =35.019, p<.009), household structure (χ^2 (15) =34.343, p<.003) and education (χ^2 (12) =33.722, p<.001). There was not a statistically significant relationship for tram users between gender, income, employment and customer fluctuation segment.

Overall, a high share of tram users appeared to be young, full-time workers with a degree and no children. Like the train market, there was a higher share of lost users that were women and lost users with children under 18. This might be an indication of public transport use being negatively impacted by the transition to parenthood.

6.4.3 Reasons for Customer Fluctuation Behaviour

One of the primary objectives of the customer fluctuation survey was to identify causal factors behind customer fluctuation segment behaviours. This provides a significant benefit over the use of anonymous smart card data which cannot provide either detailed demographic data or reasons for observed behaviour. The purpose of recording and analysing reasons behind customer fluctuation is to;

- Identify the key factors influencing customer fluctuation by mode.
- Identify how they differ by customer fluctuation segment
- Identify the ability of user-selected reasons to help predict or address customer fluctuation behaviour.

Four categories were identified based on a review of the literature to identify the reasons behind customer fluctuation behaviour; life event, personal factors, the service and the trip taken. Life events and personal factors are intended to reflect exogenous factors influencing travel decisions while the service and the trip taken attempt to capture endogenous factors. However, there is some overlap between categories. Further information is provided about this research approach in Chapter 3.

6.4.4 Disaggregate Reasons for Customer Fluctuation Behaviour

Reasons Influencing Decisions for New Users

New users were asked to identify the top three reasons that they chose to start using a public transport mode, within the four categories identified. Table 6.9 provides a summary of the factors influencing decisions for new users. Note that the share of responses are shown out of a total top three response selection; hence the total response share is out of a maximum of 300% (top 3 selection by 100% of respondents).

Table 6.8 – Summary of the top 3 reasons influencing travel behaviour for NEW USERS by Mode

Category	Reasons for Starting	Train	Bus	Tram	Total Sample[1]
Life Event	I changed home or work locations	25.0%	48.9%	59.3%	43.9%
	My life circumstances changed (I returned to work after having kids, I started a new job)	20.0%	28.9%	18.6%	22.0%
	I returned from holiday/ sabbatical	8.3%	6.7%	10.2%	8.5%
	Other	6.7%	4.4%	1.7%	4.3%
	Total	60.0%	88.9%	89.8%	78.7%
Personal Factors	I am trying to save money on my transport costs	31.7%	24.4%	30.5%	29.3%
	I was too sick or injured to drive or cycle	5.0%	0.0%	0.0%	1.8%
	I do not own a car or can't drive	11.7%	35.6%	32.2%	25.6%
	My car or bike was unavailable	6.7%	17.8%	0.0%	7.3%
	I believe it is important to take sustainable transport modes	16.7%	15.6%	11.9%	14.6%
	I have started a new hobby/ socialising more	3.3%	0.0%	1.7%	1.8%
	Other	5.0%	2.2%	3.4%	3.7%
	Total	80.0%	95.6%	79.7%	84.1%
The Service	I enjoy travelling by train/bus/tram	23.3%	26.7%	23.7%	24.4%
	I find catching the train/bus/tram reasonably priced	16.7%	13.3%	11.9%	14.0%
	I feel safest when catching the train/bus/tram	5.0%	4.4%	8.5%	6.1%
	I find the train/bus/tram to be less crowded than other modes	6.7%	4.4%	3.4%	4.9%
	I find the train/bus/tram to be reliable	10.0%	11.1%	16.9%	12.8%
	The train/bus/tram timetable changed to suit me better	1.7%	0.0%	0.0%	0.6%
	Other	1.7%	6.7%	0.0%	2.4%
	Total	65.0%	66.7%	64.4%	65.2%
The Trip Taken	I used the train/bus/tram when traveling to an event	30.0%	13.3%	16.9%	20.7%
	Parking is too difficult at my destination	33.3%	11.1%	25.4%	24.4%
	The train/bus/tram is the most convenient option for the main trips I was making	23.3%	13.3%	18.6%	18.9%
	I had an unpleasant experience with a different mode of transport	1.7%	6.7%	0.0%	2.4%
	I catch the train/bus/tram when the weather isn't suitable for other travel	5.0%	4.4%	3.4%	4.3%
	Other	1.7%	0.0%	1.7%	1.2%
	Total	95.0%	48.9%	66.1%	72.0%

Note: Participants were asked to select the top 3 most influential reasons they started travelling by a mode. These reasons were not ranked. Each column provides the proportion of new users that selected each reason by mode as such, the total is equal to 300%.

[1] The Total Sample is the proportion of new users from the total sample (train, bus and tram). However, it is noted that different sized samples were collected by mode and have not been weighted.

For the **Total** sample:

- Personal factors (84%) had the highest rate of response, followed by life events (79%) and the trip taken (72%) which had roughly similar response rates.
- Of the personal factors group, 'saving money' (29%) had the highest response rate, closely followed by 'I don't own/drive a car' (26%). 'starting a new hobby' and 'too sick to drive or cycle' had the lowest rate of response within this category.

- Of the life events group, ‘changing home/work locations’ (44%) dominated. The life event response 'changing home/work locations' was the most significant single factor cited by new users as the reason for changing their travel.
- Of the trip taken responses, ‘difficult parking’ (24%) and ‘using PT to travel to an event’ (21%) were the most common factors cited to undertake new trips, this was closely followed by ‘the most convenient option’.
- The service group had the lowest rate of responses, with ‘I enjoy travelling by train/bus/tram’ getting the highest rate of response (24%).

For the **Train** sample:

- The trip taken (95%) had the highest response rate and is higher than for the total sample (+23%). Personal factor responses (80%) were the second most common new user rail group, followed by the service (65%). Life event responses (60%) were substantially lower and less frequent than for the total sample (-19%).
- Of the trip taken group, ‘parking too difficult’ (33%) following by ‘using PT to travel to an event’ (30%) had the highest rates of response. These factors for train users were higher than for the total sample and all other modes. The train user response ‘parking too difficult’ (33%) is the most significant single factor noted in rail responses.
- Of the personal factor responses for train users, ‘trying to save money’ (32%) was the most significant response. Of the service group responses, ‘enjoy PT travel’ (23%) had the highest response rate.

For the **Bus** sample:

- Personal factors had the highest rate of response for new bus users (96%) and were higher than for the total sample. This was followed by life events (89%) and service factors (67%), with the trip taken being the least influential for new bus users (49%).
- Of the personal factors group for bus new users, ‘don’t own/drive a car’ (36%) and ‘trying to save money’ (24%) dominate. This suggests a level of captive (no choice) ridership.
- Of the life events group, ‘changed work/home location’ dominated (49%) and was the largest single factor influencing new bus user behaviour. This follows findings that people are more likely to use the bus when the service is highly convenient.
- Other influential factors included ‘changed life circumstances’ (29%) and ‘'I enjoy travelling by bus' (27%).

For the **Tram** sample:

- Life event responses were most common at almost one-third of responses (90%). This was followed by a high rate of response for personal factors (80%). Both the trip taken and service groups had similar rates of response (approx. 65%).
- Of the life event group, ‘change work/home location’ had the highest rate of response (59%) and the largest of all individual reasons for tram new users.
- Of the personal factors tram new user group, ‘don’t own/drive car’ (32%) and ‘saving money’ (31%) dominated. This might be an indication that tram users are dominated by those making lifestyle choices requiring public transport travel, such as living in the inner city and avoiding car ownership.
- The other most influential individual responses, including ‘parking is too difficult’ (25%) and ‘enjoy PT’ (24%) indicating that in certain situations, public transport is the easier alternative to driving.

Reasons Influencing Decisions for Lost Users

Table 6.9 - Summary of the top 3 reasons influencing travel behaviour for LOST USERS by Mode

	Reasons for Stopping	Train	Bus	Tram	Total Sample[1]
Life Event	I went on holiday	23.1%	18.5%	36.8%	25.0%
	I stopped studying/finished school	19.2%	11.1%	21.1%	16.7%
	My life circumstances changed (e.g. I had children, retired)	23.1%	18.5%	42.1%	26.4%
	Other	3.8%	11.1%	5.3%	6.9%
	Total	69.2%	59.3%	105.3%	75.0%
Personal Factors	I no longer needed to make trips with the train/bus/tram	65.4%	33.3%	57.9%	51.4%
	I was sick or injured and couldn't catch the train/bus/tram	7.7%	22.2%	0.0%	11.1%
	Catching the train/bus/tram was too difficult with children	7.7%	14.8%	5.3%	9.7%
	I bought a car	7.7%	33.3%	5.3%	16.7%
	I don't like the train/bus/tram	23.1%	3.7%	21.1%	15.3%
	Other	3.8%	14.8%	5.3%	8.3%
	Total	115.4%	107.4%	94.7%	106.9%
The Service	Catching the train/bus/tram takes too long	23.1%	29.6%	10.5%	22.2%
	Catching the train/bus/tram was too unreliable	7.7%	11.1%	0.0%	6.9%
	The train/bus/tram route I was using changed or stopped	0.0%	11.1%	5.3%	5.6%
	I felt uncomfortable on the train/bus/tram I was using	7.7%	11.1%	5.3%	8.3%
	The train/bus/tram wasn't available at the times I wanted to travel	3.8%	0.0%	0.0%	1.4%
	The train/bus/tram was too crowded	11.5%	3.7%	10.5%	8.3%
	Catching the train/bus/tram got too expensive	7.7%	0.0%	10.5%	5.6%
	Other	3.8%	3.7%	0.0%	2.8%
	Total	65.4%	70.4%	42.1%	61.1%
The Trip Taken	I travel at night and didn't feel safe	3.8%	18.5%	0.0%	8.3%
	The train/bus/tram wasn't needed as part of my journey anymore	34.6%	11.1%	36.8%	26.4%
	I found it easier to drive to my destination	11.5%	29.6%	21.1%	20.8%
	Other	0.0%	3.7%	0.0%	1.4%
	Total	50.0%	63.0%	57.9%	56.9%

Note: Participants were asked to select the top 3 most influential reasons in response to survey Questions 13/14/15. Reasons were not ranked. Each column provides the proportion of lost users that chose each reason by mode. The total is equal to 300%.
[1] The Total Sample is the proportion of lost users from the total sample. However, it is noted that as different sized samples were collected by mode and have not been weighted, this does not provide a fair representation of the public transport market share.

Table 6.9 provides the factors influencing travel choices for LOST Users. Core conclusions are provided below.

For the T**otal** sample:

- Personal factors (106%) was the leading category, followed by life events (75%), service factors (61%) and the trip taken (57%).
- Over one-third of the total possible responses were personal factors. Driven by over half of the respondents in this group selecting 'I no longer needed to make trips with the train/bus/tram' (59%). The next biggest reasons for being a lost user in this category were 'I bought a car' and 'I don't like the train/bus/tram' (15 – 16% each). This might indicate that users were catching public transport due to circumstances and would choose alternatives when available. What is not clear is whether users chose to travel by car exclusively or to use a different mode of public transport.
- The life event, service and trip taken groups all had a similar share of responses. The individual factors leading for lost users were 'my life circumstances changed' (26%), 'the train/bus/tram wasn't needed as part of my journey anymore' (26%), 'I went on holiday' (25%), 'catching the train/bus/tram takes too long' (22%) and 'I found it easier to drive to my destination' (21%).

For the **Train** sample:

- Personal factors (115%) were the most common factor group lost train users, followed by life events (69%), the service (65%) and the trip taken group (50%).
- Of the personal factors group, 'I no longer needed to make trips' (65%) was the most commonly selected followed by 'I don't like the train' (23%).
- For train users, there were also high responses to individual reasons in other groups. These included; 'I went on holiday' (23%) and 'my life circumstances changed' (23%), 'catching the train takes too long' (23%) and 'the train wasn't needed as part of my journey anymore' (35%).

For the **Bus** sample:

- Personal factors were again the most common (107%) for lost bus users, followed by the service group (70%), the trip taken (63%) and life events (59%). Lost bus users had the lowest proportion of users that selected life events as a factor for leaving (16%) than for the total sample.
- Of the personal factors group, 'I no longer needed to make trips with the bus' (33%) and 'I bought a car' (33%) were the dominant responses. 'I bought a car' was significantly higher for lost bus users than for the train, tram or total market (+26% compared to train and +28% compared to tram). 'I was sick or injured and couldn't catch the bus' (22%) was also higher than for either train or tram and the total sample. This may indicate that bus markets may have a higher share of captive public transport users, as well as respondents living further outside of well-serviced inner-city areas.

- Of the service group, ‘catching the bus takes too long’ dominated (30%). This factor was followed by ‘catching the bus was too unreliable’, ‘the bus route I was using was changed or stopped’ and ‘I felt uncomfortable on the bus I was using’ which each received an 11% share of responses.
- For the trip taken group, ‘I found it easier to drive to my destination’ (30%) was the most common and higher for the bus than for other modes. The bus also had the highest response rate for 'I travel at night and didn't feel safe' (19%).

For the **Tram** sample:

- For the tram, life events were the most common reason for lost users (105%); this was the only mode where this was the case. Personal factors were the second most common (95%), followed by the trip taken (58%) and the service (42%)
- The primacy of life events related to lost tram users were ‘life circumstances changed’ (42%), ‘I went on holiday’ (37%) and ‘I stopped studying finished school’ (21%).
- Of the personal factors group, lost tram users predominantly selected ‘I no longer needed to make trips with the tram’ (58%) and ‘I don’t like the tram’ (21%).
- For the trip taken group, ‘the tram wasn’t needed as part of journey anymore’ (37%) and ‘I found it easier to drive to my destination (21%) had the highest response rates.
- There were no service impacts that had a high share of influence on lost tram users.

Reasons Influencing Decisions for Returning Users (Pausing and Returning)

Returning users were asked the set of questions for both stopping travel (pausing) and starting travel (returning).

Table 6.10 - Summary of Proportion of Returning Users by Mode that identified the top 3 reasons they paused travel

	Reasons for Pausing Travel[1]	Train	Bus	Tram	Total Sample[2]
Life Event	I went on holiday	10.7%	13.6%	22.8%	15.5%
	I stopped studying/finished school	8.2%	9.1%	6.5%	7.8%
	My life circumstances changed (e.g. I had children, retired)	14.8%	13.6%	12.0%	13.6%
	Other	5.7%	0.0%	15.2%	8.1%
	Total	39.3%	36.4%	56.5%	45.0%
Personal Factors	I no longer needed to make trips with the train/bus/tram	54.1%	40.9%	55.4%	52.3%
	I was sick or injured and couldn't catch the train/bus/tram	1.6%	9.1%	3.3%	3.5%
	Catching the train/bus/tram was too difficult with children	9.8%	11.4%	4.3%	8.1%
	I bought a car	12.3%	18.2%	10.9%	12.8%
	I don't like the train/bus/tram	5.7%	9.1%	6.5%	6.6%
	Other	7.4%	2.3%	6.5%	6.2%
	Total	91.0%	90.9%	87.0%	89.5%
The Service	Catching the train/bus/tram takes too long	26.2%	40.9%	21.7%	27.1%
	Catching the train/bus/tram was too unreliable	9.0%	15.9%	4.3%	8.5%
	The train/bus/tram route I was using changed or stopped	2.5%	0.0%	1.1%	1.6%
	I felt uncomfortable on the train/bus/tram I was using (old bus, untidy, poor seats)	6.6%	6.8%	2.2%	5.0%
	The train/bus/tram wasn't available at the times I wanted to travel	12.3%	27.3%	12.0%	14.7%
	The train/bus/tram was too crowded	11.5%	9.1%	16.3%	12.8%
	Catching the train/bus/tram got too expensive	7.4%	6.8%	7.6%	7.4%
	Other	6.6%	0.0%	4.3%	4.7%
	Total	82.0%	106.8%	69.6%	81.8%
The Trip Taken	I travel at night and didn't feel safe	3.3%	9.1%	8.7%	6.2%
	The train/bus/tram wasn't needed as part of my journey anymore	36.1%	31.8%	44.6%	38.4%
	I found it easier to drive to my destination	43.4%	25.0%	30.4%	35.7%
	Other	4.9%	0.0%	3.3%	3.5%
	Total	87.7%	65.9%	87.0%	83.7%

Note: Participants were asked to select the top 3 most influential reasons. However, these reasons were not ranked. Each column provides the proportion of new users that chose each reason by mode as such the total is equal to 300%.
[1] Returning users were asked two sets of questions, reasons why they paused travel and reasons why they returned, as such, returning users are represented across table 10 and table 11.
[2] The Total Sample is the proportion of new users from the total sample. However, it is noted that as different sized samples were collected by mode and have not been weighted, this does not provide a fair representation of the public transport market share.

Table 6.10 provides the factors influencing travel choices for returning users. This table provides the answers given by returning users for why they temporarily paused travel. The reasons provided by the survey sample indicate:

For the **Total** sample:

- Personal factors (90%) had the highest response, but this was only marginally higher than the trip taken (84%) and the service (82%). Life events were the least influential reason for returning users to pause travel (45%).
- Of the three dominating groups, the highest individual factors included 'I no longer needed to take trips with the train/bus/tram' (52%), 'the train/bus/tram wasn't needed as part of my journey anymore' (38%), 'I found it easier to drive to my destination (36%) and 'catching the train/bus/tram takes too long' (27%).
- Overall, these responses indicate that for returning users looking to pause travel are primarily influenced by changing travel needs and convenience.

For the **Train** sample:

- Personal factors (91%) were again the most common closely followed by the trip taken (87%) and the service (82%).
- Of individual factors; 'I no longer needed to make trips' (54%) was the most commonly selected followed by 'I found it easier to drive to my destination (43%), 'the train wasn't needed as part of my journey anymore (36%) and 'bought a car' (23%).
- For train users, there were also high responses to the trip taken, including 'I found it easier to drive to my destination (43%). This reason was the highest for returning users across all modes. Contrary to the findings of reasons influencing permanently lost users, where driving had the least impact on train users.

For the **Bus** sample:

- The Bus was the only mode were service factors were the most selected by returning users deciding to pause travel. The prominence of service factors was driven by the response 'catching the bus takes too long' (41%), 'the bus wasn't available at the time I wanted to travel' (27%) and 'catching the bus was too unreliable (16%). These findings suggest a perceived lack of efficiency and convenience for bus travel. However, the bus did receive the lowest number of responses for 'the bus was too crowded' compared to other modes.
- The service group was followed by personal factors (91%), the trip taken (66%) and life events (37%).
- Of individual factors, the most selected were; 'I no longer needed to make trips with the bus' (41%), 'the bus wasn't needed as part of my journey anymore' (32%) and 'I found it easier to drive to my destination (25%). The bus also had the highest response rate to 'I bought a car' (19%) compared to other modes.

For the **Tram** sample:

- For the Tram, personal factors (87%) and the trip taken (87%) were tied as the main reasons influencing decisions to pause tram use.
- Of the personal factors group, returning tram users most frequently selected 'I no longer needed to make trips with the tram' (55%).
- For the trip taken group, 'the tram wasn't needed as part of journey anymore' (45%) and 'I found it easier to drive to my destination' (30%) were the leading options.
- Of the remaining factor groups, 'catching the tram takes too long' (22%), 'I went on holiday' (23%), and 'the tram was too crowded' (16%) were selected by the highest proportion of users.

Table 6.11 provides the factors influencing travel choices for returning users. This table provides the answers given by returning users for why they resumed or un-paused travel.

Table 6.11 - Summary of Proportion of Returning Users by Mode that identified the top 3 reasons they returned to a mode (recommenced travelling)

	Reasons for Returning	Train	Bus	Tram	Total Sample[2]
Life Event	I changed home or work locations	12.3%	25.0%	16.3%	15.9%
	My life circumstances changed (I returned to work after having kids, I started a new job)	4.9%	2.3%	3.3%	3.9%
	I returned from holiday/ sabbatical	1.6%	9.1%	12.0%	6.6%
	Other	7.4%	4.5%	9.8%	7.8%
	Total	26.2%	40.9%	41.3%	34.1%
Personal Factors	I am trying to save money on my transport costs	20.5%	20.5%	17.4%	19.4%
	I was too sick or injured to drive or cycle	1.6%	6.8%	6.5%	4.3%
	I do not own a car or can't drive	2.5%	9.1%	4.3%	4.3%
	My car or bike was unavailable	8.2%	31.8%	12.0%	13.6%
	I believe it is important to take sustainable transport modes	9.0%	6.8%	15.2%	10.9%
	I have started a new hobby/ socialising more	4.9%	2.3%	5.4%	4.7%
	Other	6.6%	2.3%	5.4%	5.4%
	Total	53.3%	79.5%	66.3%	62.4%
The Service	I enjoy traveling by train/bus/tram	19.7%	11.4%	19.6%	18.2%
	I find catching the train/bus/tram reasonably priced	23.8%	22.7%	19.6%	22.1%
	I feel safest when catching the train/bus/tram	0.8%	6.8%	3.3%	2.7%
	I find the train/bus/tram to be less crowded than other modes	2.5%	9.1%	3.3%	3.9%
	I find the train/bus/tram to be reliable	15.6%	9.1%	14.1%	14.0%
	The train/bus/tram timetable changed to suit me better	3.3%	2.3%	2.2%	2.7%
	Other	1.6%	2.3%	3.3%	2.3%
	Total	67.2%	63.6%	65.2%	65.9%
The Trip Taken	I used the train/bus/tram when traveling to an event	52.5%	34.1%	42.4%	45.7%
	Parking is too difficult at my destination	52.5%	43.2%	42.4%	47.3%
	The train/bus/tram is the most convenient option for the main trips I was making	45.9%	29.5%	34.8%	39.1%
	I had an unpleasant experience with a different mode of transport	1.6%	0.0%	0.0%	0.8%
	I catch the train/bus/tram when the weather isn't suitable for other travel	0.8%	9.1%	6.5%	4.3%
	Other	0.0%	0.0%	1.1%	0.4%
	Total	153.3%	115.9%	127.2%	137.6%

Note: Participants were asked to select the top 3 most influential reasons. However, these reasons were not ranked. Each column provides the proportion of new users that chose each reason by mode as such, the total is equal to 300%.
[2] The Total Sample is the proportion of new users from the total sample. However, it is noted that as different sized samples were collected by mode and have not been weighted, this does not provide a fair representation of the public transport market share.

The reasons provided by the survey sample indicate that returning users recommence travel for:

For the **Total** sample:

- The trip taken was the most selected group for recommenced travel (138%) This was followed by the service (66%), personal factors (66%) and life events (34%).
- For the trip taken group, high response rates were led by ‘parking is too difficult at my destination’ (47%), ‘I used the train/bus/tram when travelling to an event’ (46%) and ‘the train/bus/tram is the most convenient option for the main trips I was making’ (39%).
- The other individual reasons with higher response rates included ‘I am trying to save money on my transport costs (19%)’ and ‘I enjoy travelling by train/bus/tram' (18%).

For the **Train** sample:

- The trip taken was the most significant group for recommencing travel (153%), accounting for over half of all possible responses.
- Responses for the trip taken group included 'I use the train when travelling to an event' (53%), ‘parking is too difficult at my destination’ (53%) and ‘the train is the most convenient option for the main trips I was making’ (46%). The train had the highest rate of response for trip based factors across all modes.
- For the remaining individual factors, there were high response rates for ‘I find catching the train reasonably priced’ (24%), ‘I am trying to save money on my transport costs’ (21%) and ‘I enjoy travelling by train' (20%). This result might indicate that personal values, such as saving money, can appeal to users returning to train travel.

For the **Bus** sample:

- Responses for the bus were also dominated by the trip taken group (116%). This was led by those that identified 'parking was too difficult at my destination’ (53%) followed by those that selected ‘I used the bus when travelling to an event' (34%) and ‘the bus was the most convenient option for the main trips I was making (30%).
- For personal factors, there were high response rates for ‘my car or bike was unavailable’ (32%). The proportion of returning bus users that identified they only returned because their car or bike was unavailable was much higher than for the train (8%) or tram (12%).
- Of the remaining individual factors, the highest response rates were for ‘changed home or work locations’ (25%) and ‘I find catching the bus reasonably priced’ (23%).

For the **Tram** sample:

- The trip taken had the highest rate of responses (127%) followed by personal factors (66%), the service (65%) and life events (41%). 'I used the tram when travelling to an event' and 'parking is too difficult at my destination' were tied as the most prominent reasons within this group (42%). 'The tram was the most convenient option for the main trips I was taking' (35%) was the next most selected option.
- Of the remaining individual factors; the frequently selected were 'I enjoy travelling by tram', 'I find catching the tram reasonably priced' (both 19%), 'I am trying to save money' (17%) and 'I believe it is important to take sustainable transport modes' (15%).

Reasons for Retention

Table 6.12 provides the factors influencing travel choices for RETAINED Users. The reasons provided by the survey sample indicate:

Table 6.12 Summary of Proportion of Retained Users by Mode that identified the top 3 reasons they continued to use mode

	Reasons for Continuing	Train	Bus	Tram	Total Sample[1]
Life Event	I changed home or work locations	19.8%	26.2%	25.0%	23.3%
	I started studying	8.2%	10.3%	5.8%	7.8%
	My circumstances haven't changed (e.g. I take my kids to school on the bus)	9.7%	8.7%	8.2%	8.9%
	Other	3.9%	6.3%	2.4%	3.9%
	Total	41.5%	51.6%	41.3%	43.8%
Personal Factors	I believe it is important to take sustainable transport modes	27.1%	15.1%	26.0%	23.8%
	I don't have a car or can't drive	11.1%	51.6%	32.2%	28.7%
	I am trying to save on my transport expenses	23.2%	12.7%	17.8%	18.7%
	I was too sick or injured to drive	0.5%	1.6%	1.0%	0.9%
	I regularly use all modes of public transport	18.4%	26.2%	20.7%	21.1%
	Other	2.9%	0.0%	2.4%	2.0%
	Total	83.1%	107.1%	100.0%	95.2%
The Service	The train/bus/tram is the most convenient option for me	61.4%	54.8%	59.1%	59.0%
	I enjoy catching the train/bus/tram	13.0%	5.6%	9.1%	9.8%
	I feel comfortable catching the train/bus/tram	15.9%	8.7%	8.7%	11.5%
	I like that the train/bus/tram is not too crowded	2.9%	6.3%	3.4%	3.9%
	I find the train/bus/tram reliable	9.7%	6.3%	12.5%	10.0%
	I think the train/bus/tram provides a good service	6.8%	5.6%	8.7%	7.2%
	Other	1.4%	0.8%	0.5%	0.9%
	Total	111.1%	88.1%	101.9%	102.2%
The Trip Taken	I routinely make the same trip and like knowing what to do	24.2%	21.4%	16.8%	20.7%
	Parking is difficult at my regular destination	36.7%	13.5%	24.0%	26.4%
	I get the train/bus/tram as just one part of my regular journey	2.4%	17.5%	15.4%	10.9%
	Other	1.0%	0.8%	0.5%	0.7%
	Total	64.3%	53.2%	56.7%	58.8%

Note: Participants were asked to select the top 3 most influential reasons. However, these reasons were not ranked. Each column provides the proportion of new users that selected each reason by mode as such, the total is equal to 300%.

[1] The Total Sample is the proportion of new users from the total sample. However, it is noted that as different sized samples were collected by mode and have not been weighted, this does not provide a fair representation of the public transport market share.

For the **Total** sample:

- The service was the most significant factor group affecting retained users to continue travelling (102%), followed by personal factors (95%), the trip taken (59%) and life events (44%).
- The service group was led by the high response rate to 'the train/bus/tram is the most convenient option for me' (60%). No other service factors had a significantly high level of response as an individual factor.
- Other influential individual factors included; 'I don't have a car or can't drive' (29%) and 'parking is too difficult at my regular destination' (26%).

For the **Train** sample:

- The service was the most influential group for retained train users, selected by over a third of total responses (111%). As for the total sample, this was driven by participants that chose 'catching the train is the most convenient option for me' (61%).
- The service was followed by personal factors (83%), the trip taken (64%) and life events (42%).
- The leading individual factors were 'parking is difficult at my regular destination' (37%), I believe it is important to take sustainable transport' (27%), 'I routinely make the same trip and like knowing what to do' (24%) and 'I am trying to save' (23%).

For the **Bus** sample:

- For retained bus users, the most influential group was Personal Factors, accounting for over a third of possible responses (107%). Followed by the Service (88%) with The Trip Taken and Life Events both totalled at around 50% of responses.
- Personal Factors were led by a high proportion of respondents that selected 'I don't have a car or can't drive' (52%). Followed by those that chose 'I regularly use all modes of public transport (26%). Retained bus users had the lowest proportion of retained users that identified 'I believe it is important to take sustainable transport modes' (15.1%) which speaks to the likelihood that bus markets have the highest proportion of captive users without choices (Jacques et al., 2013).
- The remaining individual factors influencing bus users include 'the bus is the most convenient option for me' (55%), 'I routinely make the same trip and like knowing what to do' (21%) and ''I changed home or work locations' (26%).

For the **Tram** sample:

- Service Factors were the dominant group at a third of retained tram responses (102%) although this was tied with Personal Factors (100%).
- Service Factor responses were dominated by those retained tram users that selected, ‘the tram is the most convenient option for me’ (59%) whereas the Personal Factors group was dominated by those that selected ‘I don’t have a car or can’t drive’ (32%).
- The tram also had higher responses to personal factor items ‘I believe it is important to take sustainable transport modes’ (26%) and ‘I regularly use all modes of public transport’ (20%). This might indicate that tram users are more likely to make a personal choice to use public transport. This is likely due to neighbourhood and location factors as trams are in inner-city areas with a range of transport options.
- Of the remaining individual factors, the next most selected were ‘parking is difficult at my regular destination’ (24%) and ‘I changed home or work locations’ (25%).

6.4.5 Summary of Reasons Influencing Customer Fluctuation Behaviour

Overall the survey identified that the reasons influencing customer behaviour varied by both customer fluctuation segment and by mode. The key findings were:

- A complex mix of factors influences new user behaviour. Personal Factors like 'saving money' and 'don't have a car/ can't drive' were the most influential overall though Life Events like 'changed home/work locations' are also important. New use of Rail was most influenced by trip taken factors such as 'Parking difficulties' and 'travelling to an event'. New use of bus was influenced more by Personal factors, notably 'don't own/drive a car' and Life Event 'changed home/work location' is the most common new bus user influence.
- Lost user behaviour is predominantly influenced by Personal Factors like 'I no longer needed to make trips with the train/bus/tram'. This was followed by Life Events, which was driven by those that selected 'I went on holiday' and 'My life circumstances changed'. Service Factors were the third most selected, led by the proportion that selected 'catching the train/bus/tram takes too long. This evidence suggests that if a more convenient option becomes available, people will defect from public transport use. Results also suggest that improving service factors were likely to have a limited impact on stopping customer defections.
- Returning users that paused travel showed similar reasons to lost users and were overall driven by Personal Factors; the Trip Taken and Service factor group was also important. Again, the core reasons for pausing travel were around no longer needing to use public

transport. However, the reasons varied significantly for bus users compared to other modes, with service factors, the primary group influencing travel decisions for buses.

- Returning users that recommenced travel; the Trip Taken was the leading causal group for all modes. All modes had some variation in the proportion of responses between 'I used the train/bus/tram when travelling to an event', 'parking is too difficult at my destination', and 'the train/bus/tram was the most convenient option'. These responses indicate that a special event (e.g. a sporting event) or location-based travel (to the CBD) was a key influence for returning users to recommence travel.
- Retained users for the train and tram users are primarily influenced by service factors, while retained bus users are mainly influenced by personal factors notably not owning a car or being able to drive. Retained users for all modes are primarily influenced by the convenience of public transport options.

6.5 Discussion and Key Conclusions

This chapter addresses two of the key research objectives for this research. The first is to explore approaches for the measurement of market change segments and the second, to explore the behavioural factors influencing market change segments. This chapter found several interesting results when using primary survey data for the measurement of customer fluctuation. The rates of customer fluctuation were substantially different when using this data sample compared to the smart data approach; retained users were identified as the largest segment of public transport users across all modes. This provides a very different picture of how public transport markets operate when prepared to smart card data.

As demographic data was also available, this measurement approach also sought to identify any significant relationships between customer fluctuation segment and socio-demographic information collected. This found statistically significant relationships between customer fluctuation segment and age across all modes. Train users also had a statistically significant relationship between segment and employment as well as household structure. Bus users had a statistically significant relationship between segment and education and tram users had a significant relationship between household structure and education.

An important contribution of this chapter is the identification of the reasons influencing customer fluctuation behaviour. This found evidence of consistency amongst the reasons by customer fluctuation segment but also specific variations by mode. The similarities are briefly summarised as:

- **New users** are most influenced by changing home or work locations and not owning or having access to a car for bus and tram users. Train users were most influenced by parking being too difficult, travelling to an event or saving money on transport costs.
- **Lost users** are most impacted by changes in life circumstances for train and tram users and being able to drive to a destination for bus users
- **Retained users** are most influenced by convenience, a belief in the importance of sustainable transport modes (train and tram). Retained bus and tram users identified not owning a car as a major influence whereas train users identified it was difficult parking at their destination that drove their decision.
- **Returning users (stopping travel)** were most influenced by no longer needing to make certain trips but also train and tram users identified finding it easier to drive whereas bus users identified service based issues with availability and speed of travel.
- **Returning users (starting travel)** were consistently influenced by difficult parking and convenience when travelling to an event across all modes.

The variations by mode are largely reflective of the different contexts within which public transport services in Melbourne operate. For example, tram services focus on the inner and middle metropolitan areas and train services are designed to quickly move people into the city centre (commuting) from all areas. However, this provides important insights for how customer fluctuation must be targeted by mode as well as segment to have an effective impact. The implications of these contributions for the measurement of customer fluctuation and for policy are outlined below.

6.5.1 Implications for the Measurement of Customer Fluctuation

This thesis seeks to explore the potential for measuring and understanding customer fluctuation within public transport markets. The survey developed for this research sought to address the failure of traditional cross-sectional survey approaches to capture the temporal element of behaviour change (Saleh and Farrell, 2007, Behrens and Mistro, 2010). As found by Beige et al. (2008) the retrospective approach did allow for participants to recall their travel for the past year, however as identified in Chapter 4 there is a need for multiple years of data to better understand ongoing churn behaviours and ensure results are not reflective of seasonality. It is recommended that as identified by Kitamura and Hoorn (1987) this could be overcome by completing regular cross-sectional surveys in waves, using a different sample each time. This would allow for more accurate data, but would require additional funding to undertake and a larger sample size. However, this commitment would remain less than that required of full longitudinal studies.

The measurement of customer fluctuation using self-reported responses found high levels of customer retention and low rates of customer defections (lost users) across all modes. However, when verifying responses using the a-priori segmentation approach developed in Chapter 3: Research Approach, it was found that participants were not able to appropriately identify their own behaviour style using the descriptive statements provided. One possible cause of this is ‘social desirability bias’; where participants wanted to be identified as a user of public transport regardless of their actual frequency of use. Another possible issue is the inability of participants to consider their travel within the context of the measurement period, while not also accounting for future intent to travel. It may be more appropriate for future models to classify users based on the rules set out in the customer fluctuation model.

The results of the customer fluctuation survey also had significant differences from the findings of the smart card data analysis, as identified in Chapter 4 and Chapter 5. An analysis of the variations in results between the two measurement approaches is recommended as a next step to inform a more robust measurement process across multiple data sources. This is undertaken in Chapter 7.

The results also identified initial findings of the reasons that influence customer fluctuation behaviour. These findings, showed that the reasons influencing customer fluctuation decisions varied by both segment and mode. As an example, new and retained bus users were influenced by personal factors that suggest a strong association with captive (non-choice) users (Jacques et al., 2013). This was supported by findings of socio-demographic patterns. In contrast, train users were predominantly influenced by the trip taken, such as travelling to an event (the rail network in Melbourne covers many sporting events) or parking difficulties (the rail network provides a radial focus for Melbourne CBD where parking is more expensive).

6.5.2 Implications for Public Transport Markets/Operators

The survey approach provides two important benefits compared to the smart card approach with regards to understanding customer fluctuation – the ability to track socio-demographic information and the ability to enquire as to the reasons behind exhibited customer fluctuation behaviour. It was also noted in Chapter 3, that many public transport operators already undertake regular surveys of customer behaviour, such as the PTV Tracker Survey. As such, the modification of an existing tool may offer significant cost and time benefits when compared to introducing a stand-alone research approach.

Overall, results suggest that as anticipated, public transport markets are highly prone to fluctuation. However, this raises further questions for public transport operators as to how high levels of fluctuation might be addressed as well as needing to further understand changes in activity patterns not influenced by changes in activity patterns and/or major life events. The results in this chapter provide some contributions to address these questions.

For train users, several patterns were identified in this study, including;

- Socio-demographic data that suggested lost users may be transitioning to family, transitioning to senior employment (where users have increased salaries and time pressure and are more likely to drive) or retirement.
- New users, and returning users starting travel were predominantly influenced by travel to events (such as travel to major spectator sporting events) or travel where parking was difficult (e.g. in CBD areas). Co-incidentally lost users were most impacted by no longer needing to make trips by the train, which may indicate a switch to the private car.

These factors suggest potential strategies public transport operators might take to grow public transport markets. For example, public transport operators may benefit from targeting users that attend large spectator events for promotions to encourage increased/ongoing public transport use. Another potential strategy may be to advocate for reduced parking or an increase in parking charges at key destinations to support the use of public transport. As travel related to large events was found to be a significant driver, it may be beneficial to create a unique user type or segment (e.g. event only) that captures this type of infrequent travel. This might sit outside of the returning category, or be measured from within the returning user segment. Strategies to improve the train experience for those that have transitioned to parenthood should also be investigated.

Bus user findings suggested captivity on the bus market, with many new and retained users identifying that they did not own or have access to a car. Conversely lost bus users identified that they found it easier to drive to their destination. This was supported with the socio-demographic results which showed that higher shares of retained or returning users had a lower household income, were students and lived in single or share housing. Further, bus was one of the only modes where service factors dominated by those deciding to pause travel. Based on these patterns, it is suggested that operators consider strategies that improve or highlight the convenience of bus travel, such as advocating for bus lanes during peak hours or creating a campaign highlighting the convenience of routes.

Finally tram users demonstrated behaviours and influences falling approximately between those shown for bus and train. The tram showed similar patterns of lost users potentially associated with a transition to parenthood and new and retained users who are travelling to an event or location with difficult parking. It is suggested that a combination of strategies above could be used to improve customer attraction and retention for tram markets. Potential strategies include target users travelling to large spectator events with incentives to increase the frequency of their travel and investment in separated tramways (permanent or during peak hours) to improve the quality of travel.

These strategies are suggested to provide an example of how the insights in this chapter could be used for industry marketing practice as a method to provide new insights and investigate links between customer fluctuation groups and service modification or marketing strategies. Further research is required before any strategies are implemented, including a cluster analysis which may be used to identify further demographic and personality characteristics of customer fluctuation segments. The continued use of customer fluctuation as a measurement tool would allow for the monitoring of the impact that strategies might have on ridership overall and the rates of customer fluctuation for each mode of public transport. However, as discussed in the next chapter, there are several refinements to the concept that will be required.

6.6 Conclusions

This chapter explored the use of a cross-sectional survey as a tool to measure customer fluctuation and the factors that influence customer fluctuation behaviour. The first part of the chapter used the results of the customer fluctuation survey to measure the rate of customer fluctuation across all modes of public transport within Melbourne. This found that the survey could measure customer fluctuation, however verification was required as there were discrepancies between recorded travel patterns and participant's self-classification into customer fluctuation segments. There were also significant differences in the results using self-classification from previous research using smart card data in Chapter 4 and 5. This will be the focus of the following chapter.

This chapter reviewed the socio-demographic patterns of each customer fluctuation segment by mode. Patterns were found to be based on mode rather than across the customer fluctuation segment. Regardless, this provided new insights that were not captured through smart card measurement approaches and could be further refined.

Finally, there was an analysis of the reasons provided by participants for their customer fluctuation behaviour by mode for each segment. This was a core aim for the use of a cross-sectional survey, as it provided the only ability to ask users to explain their travel behaviour. Findings identified that different factors influenced the decisions of different segments. It was identified that improved clarity and separation between categories and reasons provided is necessary to develop a deeper understanding of the motivations influencing customer fluctuation behaviour. This was particularly the case for lost users and returning users that paused travel.

As this research is the first exploration of the customer fluctuation concept, it might be expected that problems such as those with self-reporting of segments would be encountered. The next chapter, Chapter 7, compares the two research approaches adopted to measure customer fluctuation and seeks to develop a refined approach that could be used in future research.

Part C: Concept Review and Conclusions

Chapter 7: Comparison of Measurement Approaches

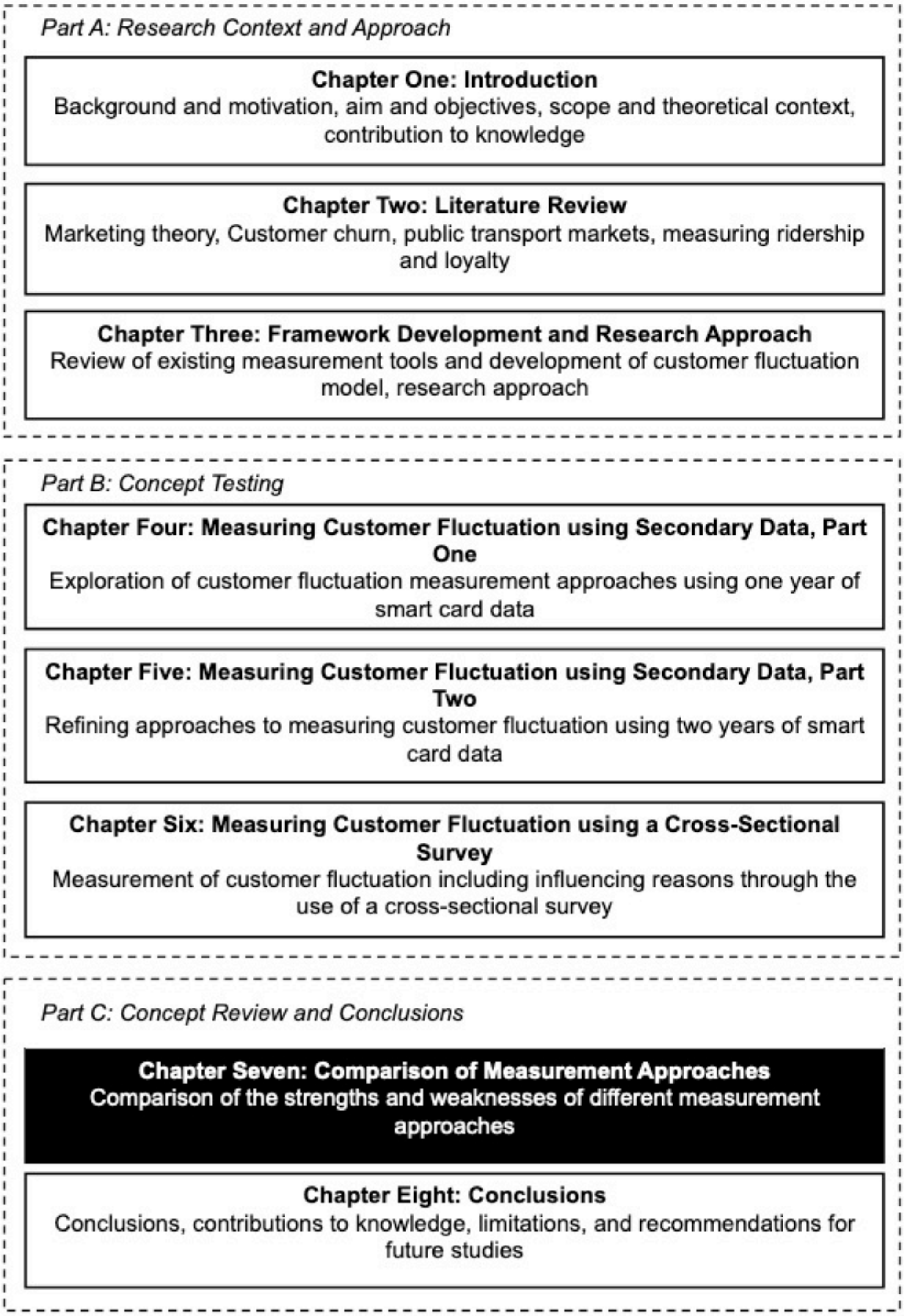

Figure 7.1 - Position of Chapter 7 in the thesis structure

7.1 Introduction

The previous chapters present methods for the measurement of customer fluctuation using smart card data (Chapters 4 and 5) and primary survey data (Chapter 6). The approach of developing different methods using different information has been utilised as a form of triangulation. Triangulation is grounded in the principle that any isolated method of obtaining data has weaknesses (Connidis, 1983) and more robust inductive reasoning can be achieved by converging different forms of information about the same phenomenon (Denzin, 1978, Jick, 1979). However, the two measurement approaches result in two distinctly different sets of findings. This chapter investigates the differences between the two measurement approaches for customer fluctuation and proposes an updated measurement approach.

The chapter is structured as follows; first, a comparison of customer fluctuation results is made with previously published evidence in the field. The nature of the differences between data sources is then provided, followed by the potential causes for the difference. The chapter then presents an updated approach incorporating the key strengths of the two methods using smart card and survey data.

7.2 Research Aims and Objectives

This chapter responds to the second aim of this research; this aim is restated as follows:

II. To explore the potential of this approach as a means of improving public transport market change analysis

This chapter responds to Research Objective 5 in particular, this objective is:

RO5. To assess the potential of the new approach for application in the industry

7.3 Comparison to Previously Reported Results

Mason et al. (2011) studied customer churn in public transport and the findings are used to assess customer fluctuation results of this research. The comparison has some limitations which include:

- The findings in Mason et al are for Long Distance British Rail Markets only and do not cover other modes
- The Mason et al study is of a different market context and thus is likely to have different results

- The Mason et al study did not evaluate the impact of returning users on the market.
- The Mason et al study did not account for the entire market but was divided into only commuter and leisure market areas. Commuter data is adopted to make a comparison with estimates from the survey/smart card as these were estimated with a greater level of accuracy in Mason et al research.

The results of the verified survey data for the train are re-segmented in accordance with the definitions identified in the Mason et al. (2011) study. An approximate comparison, using the available details of the method proposed by Mason et al. is provided in Table 7.1.

Table 7.1 - Comparison of Primary and Secondary Data to Mason et al (2011) Results

	Composition of Customer Sample				Churn Measures		
Data Source	**All**	**Loyal**	**New**	**Lapsed**	**Customer Retention**	**Market Entry**	**Market Exit**
Mason et al (2011) Data							
Total British Rail Markets Survey[1]	100%	63%	19%	18%	77%	23%	23%
Project Data							
Melbourne Train Market Survey[2*]	100%	57%	13%	30%	57%	26%	30%
Melbourne Smart Card Train Users[3*]	100%	18%	11%	71%	18%	47%	71%

Note: [1] Mason et al, measure train travel between 2009 – 2011 via a survey
[2]The Customer fluctuation survey (verified results) measures train travel between 2017 - 2018
[3]The smart card data utilises measures train travel between 2017 – 2018
* Data has been scaled using an approximation of the segmentation process undertaken by Mason et al (2011)

This comparison highlights a relative consistency between the survey method results and those obtained in Mason et al. (2011). Smart card method results in a substantially higher share of lapsed users (market exit) and a lower share of loyal or customer retention segments. As the Mason et al and survey method results are close, this provides some encouragement that the results achieved are comparable with our benchmark study. Regardless, it is difficult to discern symptomatic causes of variation in measurement approaches given the limitations noted about comparisons between the two data sets. The following discusses the differences between the survey results and smart card findings of this analysis and possible causes.

7.4 Differences between Smart Card and Survey Results

As shown in Table 7.2, the most substantial differences are that survey results had a much larger share of retained users compared to smart card methods. Also, lost users are a relatively lower share in the survey approach compared to the smart card approach.

Table 7.2 - Visual Representation of the Data Analysis Problem when Measuring Customer Fluctuation

	Smart Card (2 Year with Verification Buffer)		Survey (Verified)
New		→	+
Lost		→	- -
Retained		→	++
Returning		→	-

Note: The visual depiction shows potential gains in users in blue, with + signs indicating the general strength of the impact. Negative impacts are shown in orange, with– signs indicating the strength of the impact.

Key:		**Base results**	+	**Increasing**	-	**Decreasing**

This presents the following problem:

> **There is substantial variability in the results from secondary (smart card) and primary (survey) data sources for the lost and retained market segments.**

There are limited explorations of customer fluctuation (or churn) as applied to public transport markets in literature. These variations cannot be verified against previous work. Therefore, further exploration is required to better understand why these variations have occurred and whether the methods used can be adjusted to mitigate these effects.

7.5 Analysis of Potential Causes

Table 7.3 provides a more detailed assessment of the differences in findings between the two methods, highlighting differences in market segment estimates by public transport mode. In this table, the survey results that have decreased the share of users in a segment substantially from the smart card results are shaded orange. Conversely, where the verified survey shares are higher than those found through the smart card data analysis, the change has been shaded in blue. The relative scale of difference is also highlighted.

This assessment highlights that the patterns of differences in estimates are universally consistent across modes, as all modes showed the same pattern of change between the smart card and survey results.

Table 7.3 - Detailed Variations in Customer Fluctuation Segment Results by Mode and Method

		Smart Card 2-year (with buffer)	Survey (verified)	Change
Train	**New**	3.6%	8.1%	+
	Lost	50.0%	16.8%	--
	Retained	3.8%	48.4%	++
	Returning	42.6%	26.8%	-
Bus	**New**	3.6%	10.0%	+
	Lost	62.4%	16.5%	--
	Retained	2.1%	50.5%	++
	Returning	32.0%	23.0%	-
Tram	**New**	5.3%	10.9%	+
	Lost	56.5%	7.2%	--
	Retained	2.6%	56.8%	++
	Returning	35.5%	25.1%	-

Key: + Increase - Decrease

Notes: This comparison focuses on the two preferred measurement modes – the smart card 2-year data with verification buffer and verified survey results.
The number of +/- indicates the size of the difference observed

Table 7.3 also illustrates the scale of differences in market segments shares between smart and survey methods. It identifies that the differences affect the customer fluctuation segments in the following order of scale (most to least difference);

- Lost Users
- Retained Users
- Returning Users
- New Users

Estimates of the share of the lost user segment are substantially lower from the survey approach compared to the smart card approach; for tram (-49%), bus (-46%), and train (-33%). This is a significant concern as the retention of lost users is a primary influence on the measurement of customer churn and of great concern to public transport operators. The second largest variance is in the retained user segment, where share estimates are larger using the survey method population compared to the smart card method. This difference is in the order of +45% for the train, +48% for the bus and +54.2% for the tram. The scale of differences is much less for the following two segments. Returning users have a smaller proportion of users for the survey method when compared to the smart card method. This is -16% for the train, -9% for the bus and 10% for the tram. The difference in the proportion of new users is the least significant difference with share estimates higher for the survey method compared to the smart car method at +5% for the train, +6% for the bus, and 6% for the tram.

The following discussion explores the possible causes of these differences. The most significant potential causes impacting on the survey include the following:

- **A survey response bias**: Those that participate in a survey on public transport issues are more likely to be active public transport users. This is a common issue with surveys, in this instance, it is likely to exaggerate the number of retained and returning users while reducing the proportion of those that no longer use public transport (lost users). The use of a market research company to collect survey responses attempted to limit this bias by offering a small participation fee and setting a quota for non-public transport user responses.
- **Social desirability bias**: There was evidence in the survey results of a social desirability bias that skewed results in favor of retained users. This is a bias where participants thought of themselves as being pro-public transport use despite the actual behaviour patterns they recorded. The bias was evidenced by the tendency of participants to incorrectly self-categorise themselves as a retained public transport user despite their travel patterns being more consistent with returning or lost use.
- **Survey development and design**: This survey asked users to respond only for their mainly used mode. This approach was chosen to limit the time imposition on survey respondents who may use several public transport modes. The impact of this decision is likely to emphasise retained users as it focuses on the primary mode and neglects to account for incidental or more irregular travel patterns that may occur for the same user on their secondary modes of transport.

There were also several potential causes for the data differences that could be linked to the use of smart card data. These potential causes and their likely impacts are briefly summarised as followed:

- **Smart Card Expiry**: Currently, myki cards expire after four years and there is no process for linking the expired card of an existing user to their new card. This is likely to result in the over-representation of lost users within the sample, as those with expired cards effectively disappear for the purposes of behaviour analysis.
- **User error and variance**: The use of smart cards is not consistent from individual to individual. Anecdotally it is known that many users hold multiple smartcards for ease of use or may use and replace smart cards over the measurement period.
- **Sampling Issues**: As previously identified, there is limited demographic data associated with the travel behaviour data derived from smart cards. Unlike survey data, which collected a sample of Melbourne based public transport users, smart card data cannot impose a sample frame. As such, the data for Melbourne also includes interstate and international tourists. These sporadic users are likely to inflate the proportion of lost users measured.

The potential impact of these causes on the results are illustrated in Table 7.4.

Table 7.4 – Potential Causes of Differences Between Smart Card and Survey Data and the Likely Effect on Estimate Share

	Smart Card				Survey			
	Likely Impacts on Estimates							
Potential Cause	New	Lost	Retain	Return	New	Lost	Retain	Return
Survey response bias in favour of interested public transport users						-	++	
Social Desirability bias where people want to be pro-public transport despite actual behaviour						--	+	
Survey working required users to respond for their mainly used mode							++	
Smart cards expire every four years and new smart cards are not linked to the trips of previous smart card users		++	-					
Users may hold multiple smart cards or may lose a smart card and be required to replace it		++	-					
Smart card data will also include a small share of tourist users, who will be identified as lost		+						
	Summary of Overall Differences in Estimates between Smart Card and Survey Data							
Overall differences		++	--			--	++	

Note: The visual depiction shows potential gains in users in blue, with the number of + signs indicating the general strength of the impact. Negative impacts are shown in orange, with the number of – signs indicating the strength of the impact on the data set.

The number of +/- indicates the estimate size of the difference

Key:		No impact	+	**Increase**	-	**Decrease**

Overall, the net aggregate effects of the potential causes discussed above suggest that:

- Compared to survey methods, smart card methods might have lost market segment shares which are higher; and
- Compared to smart card methods, the survey estimates of retained market segment shares might be expected to be higher.

This is a significant finding because these theorised impacts match those found in the differences between market segment share estimates from the two methods.

We therefore, conclude that data collection biases and errors in both the survey and smart card methods of measuring market segments act to create different results between the two methods. The following sections address some recommended additional areas for analysis to built our understanding of customer fluctuation (7.6) and then proposes an improved measurement approach accounting for these biases to improve the accuracy of estimates (7.7).

7.6 Additional Areas for Analysis

One of the goals for this research was to test the potential of the new approach of customer fluctuation for application in the industry. Due to the significant development required to create the new measurement concept of customer fluctuation and the time limitations of this research the analysis was required to be kept at a high-level overview of its potential. Further this was done due to the significant differences in customer fluctuation results between data sources and lack of data verification available. There are several further areas for study that would be of great interest for the further development of this concept, these include:

- Completing post-hoc cluster analysis to test the findings of the a-priori segmentation process and identify whether these groups do contain similarities. Further, a clustering analysis based on both personality factors and ridership may allow for deeper insights for practitioners and marketers (Anable, 2005).
- Disaggregate analysis of two interesting findings from the survey conducted in Chapter 6 – the indication of a social desirability bias and the impacts of pro-environmental factors on public transport use. Interesting insights might be gained by investigating the impacts of this by customer fluctuation segment.
- Use of a local expert reference group or panel to validate the overall results from both measurement approaches. This would allow them to review the results with regards to existing data sets and measurement of Melbourne markets, and would improve on the imperfect comparison with the Mason et al (2011) study.
- The application of customer fluctuation to different public transport contexts to add information about differential market change across different geographies. This would also add further knowledge to help verify the appropriate customer fluctuation measurement approach as different public transport markets may have improved data sets.

As these areas of additional research may offer an efficient measurement tool and new insights for public transport operators, helping to draw distinct conclusions on links between customer fluctuation segments and marketing approaches/service changes. It is considered that the potential of the tool warrants a further investment of time and research. The next section provides a suggested adjusted measurement approach to address the limitations encountered with the measurement approaches used in this thesis. This section has been completed without considering the results of any of the above additional areas for analysis. The additional analysis areas completed above may be completed before or after the use of an adjusted measurement approach.

7.7 An Adjusted Measurement Approach

Based on the analysis of potential causes for the data differences and the overall potential of the concept, an adjusted research approach is proposed. It is considered that the measurement approach can be improved through the integration of data sources in a hybridised approach.

Overall, it is recommended that an improved approach to measuring customer fluctuation combines both smart card data and survey data to create a more accurate and complete single method of measurement. A panel of respondents would be required that consent to having their smart card information accessed and linked to any survey information they provided. The use of a panel would ensure the ability to track longitudinal data from specific individuals while ensuring the data is broadly representative of the Melbourne population. Similar approaches have been utilised by Ji et al. (2019), Medina and Arturo (2018), Li et al. (2018b). A summary of these approaches is provided in Table 7.5.

Table 7.5 - Summary of Literature Using Mixed Smart Card and Survey Research Approach

Article	Purpose	Method	Limitations
Medina et al., 2018	Identify weekly activity patterns	Using a household travel survey (pre-existing for 1% of households) and smart card data for Singapore. Clustering of activities using DBSCAN algorithm.	Splits observations between workers and students. Activity findings are not associated with socio-demographic data
Li et al., 2018	Analysing long term travel behaviour between smart cards and survey information	Smart card data is matched with a GIS survey where users provide a stated preference. Comparison suggested that the pattern specific average usage (actual usage) comply fairly well with the proposed proxy patterns (stated usage), although there is a bias towards more recent travel.	Required manual matching between survey data and smart card transportation records. Measured increased use, decreased use and stable use. Also, does not link the findings to socio-demographic data
Ji et al., 2019	Classification and influencing factors of metro commuting patterns by combining smart card data and household travel survey data	First, the authors generate the commuting regularity rules using one day household survey data. This is clustered using the Gaussian mixture model, they classify smart card data into three commuting pattern groups, namely, classic pattern, off-peak pattern, and long-distance pattern, based on spatiotemporal characteristics. Next, they link commuters of these three groups to the household survey and apply a mixed logit regression model to determine the factors influencing commuting patterns.	Short term study that doesn't measure long term variations in ridership patterns. Also, commuting only focus. Does not account for passengers' perception/attitude/tolerance of travel conditions.

All studies presenting in Table 7.5 were approaches utilised to identify information beyond the measurement of ridership patterns or customer fluctuation. It is suggested that a simplified approach can be used to measure customer fluctuation based on the approaches utilised above.

The following sections discuss two potential approaches to create a hybrid data source for the measurement of customer fluctuation; the use of an existing smart card panel (run by PTV) or an independent longitudinal survey linked to smart card data. Several key changes should be made to these two approaches for customer fluctuation measurement:

- The use of a larger sample of participants whose smart card data is tracked across how they use all three modes available in Melbourne.
- The inclusion of a new segment or sub-segment of returning users to account for event only travel.
- The removal of any requirement for participants to self-categorisation their public transport behaviour within customer fluctuation segments
- Participants to be sent a survey on the reasons behind their travel behaviour, after they have been categorised using the a-priori segmentation approach. This will allow for participants to complete a brief survey relevant to their actual behaviour and verified customer fluctuation segment.
- The refinement of the reasons for behaviour suggestions to improve clarity and ensure categories are distinct from each other.

7.7.1 PTV Smart Card Panel

Public transport operators, such as PTV, have significant research resources dedicated to understanding public transport markets using smart card data. PTV has a panel of smart card users who have consented to have their smart card recorded alongside relevant demographic details of each user who are also willing to respond to surveys about their transport behaviour. In this case, the panel may be used to test customer fluctuation by surveying panel members about their travel behaviour over the past two years and using an associated a-priori segmentation approach on their smart card data records.

This approach would be both cost and time effective, however, there are some limitations that must be taken into consideration.

- Any incentives that panel participants receive, (free or reduced trip fare) are likely to result in travel behaviour patterns that are not necessarily typical of the market.
- It is important that those recruited onto the panel are representative of all public transport users. Otherwise, this might represent an inherent bias.

- The panel needs to be of sufficient size to maintain a degree of statistical accuracy to estimate market segment size (a minimum of 95% CI (+/- 5%)).
- Where two years is considered a prohibitive time period to run a panel, an 18-month panel may be considered to provide a 12-month measurement period with 3 months for verification at the beginning and end of the measurement period. However, this approach might limit the ability to understand seasonality based changes in ridership.

Any existing smart card panel should be assessed in terms of these limitations to determine whether an improved data source could be provided through this method. Further, this process must be ongoing to allow for repeat interviews about public participation behaviour.

7.7.2 An Independent Survey Linked to Smart Card Data

An alternative approach is to complete an independent survey that will require participants to consent to their smart card data being made identifiable and linked to their survey responses. This approach would enable retrieval of smart card data and its associated monitoring of travel as trips per month for a unique smart card ID, while also understanding the associated demographic and attitudinal details of the card user. This approach can be used to estimate customer fluctuation segments and to monitor user responses to questions such as reasons for changing travel behaviour. This approach has the advantage of addressing the issue with self-selection bias identified in the initial survey results.

An independent survey could also be conducted as a longitudinal survey or a retrospective survey dependent on time and resource limitations. A longitudinal study provides for more accurate participant recall, however, would require significant time and resource investment, as it has been found that two years is a preferred measurement period for capturing customer fluctuation. Further longitudinal studies are prone to a decline in rates of participation over time. A combination approach may be used where participants are asked to recall the reasons for their travel behaviour for the previous year and then followed for the next year and asked to identify any differences. The measurement of customer fluctuation would still rely on the a-priori measurement approach utilised throughout this thesis. A concept diagram of the proposed research approach structure is provided below as Figure 7.2

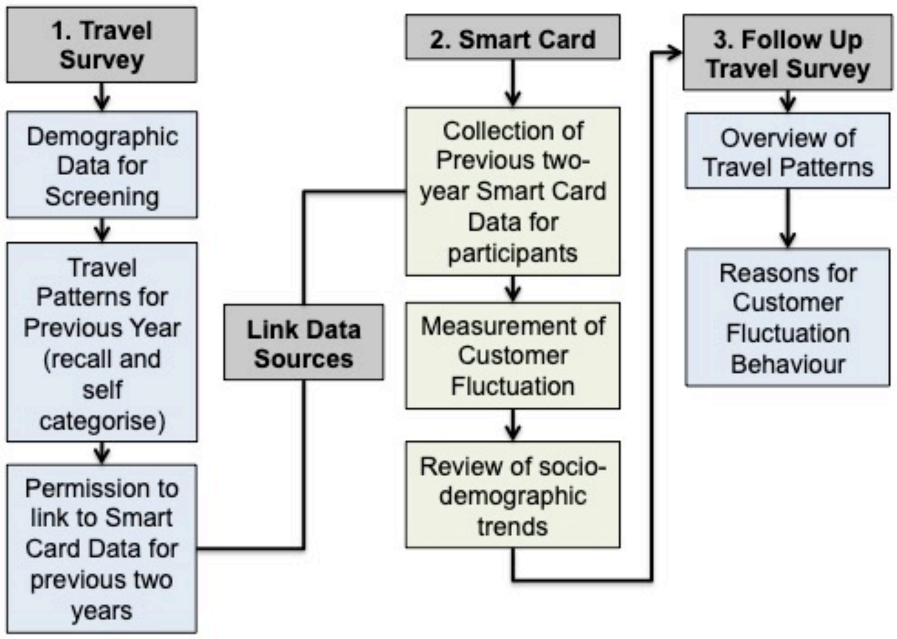

Figure 7.2 - Diagram of Proposed Improved Research Approach to measuring Customer Fluctuation

Overall, the integrated approach improves on the following limitations found in the initial approaches (Chapters 4, 5 and 6) used to measure customer fluctuation:

- The ability to combine smart card and travel survey data to produce a more accurate measurement of the actual rates of customer churn occurring.
- Allowing for the verification of customer fluctuation and identification of reasons for fluctuation behaviour via a follow-up survey.
- A mixed method to address the limitations of both smart card and survey data in measuring customer fluctuation.

There are limitations that remain with this research approach. The process will continue to over-state the appearance of stable users as it requires two or more years of public transport use. It will be important for new participants to be regularly added to this approach to ensure new and lost users are appropriately captured.

With further testing, an improved research approach using the hybrid model may be used to train a model to predict customer categories from a survey based on the travel patterns extracted from their smart card data. This would make it possible to use smart card data to categorizing users into segments with improved accuracy.

7.7.3 Anticipated Impacts of Revised Approach

The proposed changes to the research approach have been designed to address the potential causes for data differences between smart card and survey data. The anticipated impact of these changes is summarised in Table 7.6.

Table 7.6 - Summary of how the proposed approaches are likely to address potential causes for data differences

	Revised Impacts on Estimates with proposed approach							
	PTV Panel				**Independent Survey**			
Potential Cause	New	Lost	Retain	Return	New	Lost	Retain	Return
Survey response bias in favour of interested public transport users		-	+			-	+	
Social Desirability bias where people want to be pro-public transport despite actual behaviour		n/a	n/a			n/a	n/a	
Survey working required users to respond for their mainly used mode		n/a	n/a			n/a	n/a	
Smart cards expire every four years and new smart cards are not linked to the trips of previous smart card users		n/a	n/a			n/a	n/a	
Users may hold multiple smart cards or may lose a smart card and be required to replace it		+	-			+	-	
Smart card data will also include a small share of tourist users, who will be identified as lost		n/a	n/a			n/a	n/a	
Note: The visual depiction shows potential gains in users in blue, with the number of + signs indicating the general strength of the impact. Negative impacts are shown in orange, with the number of – signs indicating the strength of the impact on the data set. The number of +/- indicates the estimate size of the difference. Where the impact is expected to be removed or negligible, a n/a is given.								
Key:		Mitigated or limited impact	**+**	**Increase**	**-**	**Decrease**		

How these changes have been addressed is briefly summarised below:

- **Survey response bias:** Partially addressed by both approaches though dependent if the panel is representative of Melbourne and not public transport users in Melbourne when using the PTV panel. May have an increased ability to influence this through an independent hybrid survey. This is still likely to have some impact due to the nature of the research and a focus on public transport.
- **Social desirability bias:** Addressed by both approaches. Participants will not be asked to self-categorise their travel behaviour. Follow up surveys will be used to identify the reasons for travel.
- **Survey design and focus:** Addressed by both approaches as participants will be required to respond for their travel across all three modes.
- **Smart Card expiry:** This is addressed by both approaches as participants will be able to report whether they have an expired smart card and are required to provide an updated smart card ID.
- **User error/variance with smart cards:** Lost and replaced smart cards can be recorded by both approaches as with the above. User error or holding multiple smart cards is unlikely to be addressed through the proposed approaches.
- **Smart Card Data Sample:** Both approaches will have linked demographic profiles and information so that no tourist cards will be included through the proposed changes.

Though there will still be some limitations with the approaches proposed, this is considered to present a significant refinement of the issues identified. Importantly, a hybrid approach will provide a single measure of customer fluctuation that is subject to consistent limitations across multiple data sources. The proposed changes are considered to provide a more accurate picture of customer fluctuation.

7.8 Discussion and Conclusions

This chapter compares the Customer Fluctuation measurement approaches, including those used in this research as well as external studies. The initial exploration of measuring churn as presented in this research, found significant differences in the results when using smart card method compared to the primary survey method. These differences were most significant for the retained and lost user segments, with a significantly higher proportion of retained users estimated when using survey measurement and a lower proportion of lost users. There was also some variation in the proportion of returning and new users, though these segments did not exhibit large differences between approaches.

The analysis identified several potential causes for these differences in results. The potential causes impacted the survey via response bias, social desirability bias, the choice of 'mainly used mode' to target survey responses. However, there were also causes impacting the accuracy of smart card data, including smart card expiry dates, lost or replaced smart cards, and the inclusion of tourists and users outside of the metropolitan Melbourne market (as targeted in the survey). A review of these potential causes identified that their likely impacts on the estimated market segment size reflect the actual variation identified in results between the two measurement approaches. We conclude that measurement issues and biases in each method act to create the estimation differences found in practice.

As there are limited explorations of customer fluctuation (or churn) as applied to public transport markets in the surrounding literature, these variations cannot be verified against previous work. The work was compared to the existing research available on customer churn, as presented by Mason et al. (2011) for British railway markets. This comparison found a relative consistency between the estimates made from the survey method and those in the Mason et al. (2011) study. However, it's unclear what can be concluded from this as there are many limitations in adopting the Mason et al approach.

The above findings were used to suggest an improved research approach for the measurement of customer fluctuation within public transport markets. Based on a review of current literature using a combination of smart card and survey data to understand public transport markets, a similar approach is proposed for the measurement of customer fluctuation. The approach would involve a representative panel to undertake a longitudinal study of their individual public transport use behaviours. This approach would involve an initial survey stage, where the panel is asked about their current public transport use and are required to link their smart card to their answers to allow for travel verification. This data would then be used to track and measure customer churn in accordance with the a-priori method identified within this research. Participants would have their travel followed for a year before being asked to complete a second survey to verify the segmentation and identify reasons for their customer fluctuation behaviour. It is recommended that this approach would address the key limitations identified in chapters 4, 5 and 6 and would allow for an improved measurement of customer fluctuation.

Chapter 8: Conclusions

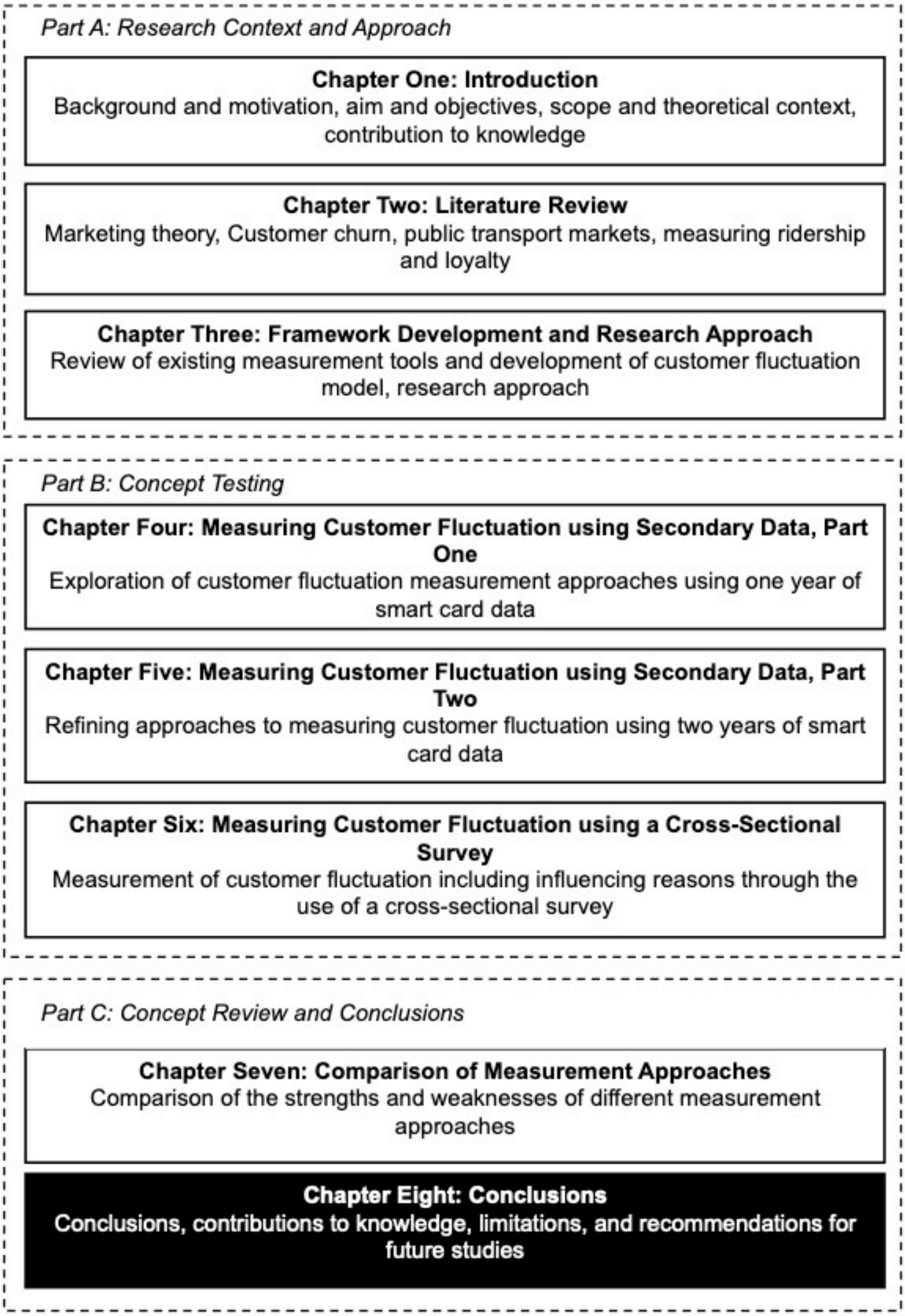

Figure 8.1 - Position of Chapter 8 in the thesis structure

8.1 Introduction

This final chapter brings together the overall findings of this thesis. This chapter reiterates the aims and research questions of this thesis, and then provides a summary of how the overall findings address the research questions and sub-questions. This is followed by a summary of contributions to knowledge resulting from the research and a critique of the research approach. The chapter then presents the recommendations for future research and the overall conclusions of this thesis.

8.2 Research Aims and Objectives

As stated in the introduction, two research aims guide this thesis:

I. To develop, measure and apply a new concept for public transport market change analysis based on segments that represent changes in ridership ('market change segments')

II. To explore the practicalities and potential of this approach as a means of improving public transport market change analysis

These aims focused on the development of a new concept which is termed Customer Fluctuation. The thesis explored the measurement, practicalities, and potential of this new approach to market change analysis.

To guide this exploration of Customer Fluctuation, five key research objectives were set:

RO1. *To understand conventional measures of market change in the fields of marketing and public transport and explore the benefits and drawbacks of these methods*

RO2. *To develop a new concept for market change analysis based on market change segments*

RO3. *To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne*

RO4. *To explore behavioural factors influencing market change segments using survey data for metro Melbourne*

RO5. *To assess the potential of the new approach for application in the industry*

8.3 Summary of Key Findings

This section provides a summary of the key findings from this thesis, organised by how they respond to the research objectives.

8.3.1 *To understand conventional measures of market change and explore the benefits and drawbacks of these methods (RO1)*

This thesis explored the intersection between marketing literature and public transport through an exploratory literature review. This focused on the marketing concept of customer churn, which measures the ‘tendency for customers to defect or cease business with a company’ (Kamakura et al., 2005). This concept is valuable as marketing theories suggest that retaining customers by limiting defections is more likely to grow markets than alternative methods focused on recruiting new customers.

Customer churn has traditionally been applied in contractual settings (e.g. phone contracts, credit cards) to measure the rate that customers defect or re-sign up for a service. The application within contractual settings is straightforward as there is a clear connection between service use and retention as well as discrete time frames for re-subscription or renewal. This makes it easy to distinguish between a lost customer and an existing customer. However, customer churn has been applied in non-contractual settings with some modifications such as defining measurement periods and thresholds for defection (Ascarza and Hardie, 2013, Buckinx and Van den Poel, 2005, Tamaddoni et al., 2010).

This exploration found that measuring the level of market churn is beneficial as it allows us to set benchmark rates of acquisition and defection within an industry or market. This information can also be used to identify what degree of observed customer defections is unusual (Riebe et al., 2014). This is useful information for service providers, as an unusually high rate of defections may be an indication of customer dissatisfaction with current service provision. It can, in some cases, also allow for comparisons to be made with competitors to identify competitive advantage or disadvantage. Furthermore, measures of churn can be used to build predictive models of customer behaviour; this can be used to predict when customers are likely to defect and target them with offers or incentives prior to the churn event (Holtrop et al., 2016, Bellman et al., 2010).

There were also several drawbacks with customer churn models identified in the literature review. The most significant drawback was that measures of customer churn typically focus on a single relationship between lost and retained customers (Kamakura et al., 2005, Tamaddoni et al., 2014, Sharp et al., 2002). This neglects two further influences on market change: the impact of new customers and cycles of customer defection and return. Taken together, four key elements of market change provide a more complete picture of market dynamics: acquisitions (new users), defections (lost users), retention (retained users), and sporadic use (returning users). To our knowledge, this thesis is the first time these four components of market dynamics have been considered together.

These four elements of market change were used to categorise studies of market change and adjacent fields within public transport markets. This found that there are a variety of approaches for measuring market change within public transport, though few focus on individual changes in ridership and even less measure all four elements of market change. A synthesis of these findings is provided in Table 8.1.

Table 8.1 – Synthesis of applicable literature against the four key elements of market change

Reference	Acquisitions	Defections	Retention	Internal Variability	Public Transport Markets
PT/CHURN					
(MASON ET AL., 2011)	Yes	Yes	Yes	Yes	Yes
(SALEH AND FARRELL, 2007)	No	No	No	Yes	Yes
PT/ NET MARKET CHANGE					
(CURRIE AND WALLIS, 2008)	Yes	Yes	No	No	Yes – Bus market focus
(CHEN ET AL., 2011B)	No	No	No	Yes – Aggregate, macro-level	Yes
PT/ LOYALTY					
(TRÉPANIER ET AL., 2012)	No*	No*	Yes	Yes	Yes
(VAN LIEROP AND EL-GENEIDY, 2016)	No	No	Yes	No	Yes
(BASS ET AL., 2011)	No	Yes	Yes	No	Yes
(TAO ET AL., 2017A)	No	No	Yes	Yes	Yes
PT/ VARIABILITY IN BEHAVIOUR					
(CSIKOS AND CURRIE, 2008)	No	No	No	Yes	Yes
(BRIAND ET AL., 2017)	No*	No*	No	Yes	Yes
(CHU, 2015)	No	No	No	Yes	Yes

*Although measured, not utilised in discussion.

Only a single study was identified that reviewed public transport markets with regards to all four elements of market change. This study, by Mason et al. (2011), used a profile survey to identify customer churn in British Railway markets over a two-year period and created a specific measurement of churn in commuter travel and a broader measurement of churn in leisure travel. This study identified that due to the changing external conditions across the two years, the underlying level of churn might be over-estimated. Though this study provided a valuable starting point for the measurement of customer fluctuation, its limited scope (railway markets only) and lack of clarity around the methodology suggest a need for further refinement.

Though the focus on customer retention is the most significant drawback of churn measurements, other drawbacks were also identified in the literature review. Brady (2014), identifies that one of the key issues with measuring churn is identifying the relevant variables for use in particular temporal elements. The temporal scale for the measurement of customer churn remains difficult to identify as it changes between industries and settings (contractual vs non-contractual). Another drawback, as identified by Tamaddoni et al. (2016), is that the definitions of churn imply that the loss of customers is a permanent state. This makes churn measurement less relevant to non-contractual service settings, such as public transport, where customers have no obligation to repurchase or re-patronise a service within a certain time frame and may stop and start using these services over time.

In summary, conventional measures of market churn tend to overlook the complete dynamics of customer usage patterns. There are also potential issues with applying a market churn measure to non-contractual settings such as public transport use. These initial findings will help guide the exploration of the remaining research objectives.

8.3.2 *To develop a new concept for market change analysis based on market change segments (RO2)*

The analysis of current approaches to market change measurement (customer churn) identified several drawbacks that could be improved upon for the application of such approaches to public transport markets. This led to the creation of a new concept for this thesis: **customer fluctuation**. The concept of customer fluctuation creates disaggregate market segments which capture the internal variability, or fluctuation, of ridership within public transport markets. We defined the concept as:

> *"A concept that seeks to measure market change over time by separating markets into disaggregate segments based on changes in ridership (for example starting, stopping or continuing to travel). This concept measures the interplay between new, lost, retained and returning customers within public transport markets."*

This concept was developed to improve upon the traditional application of customer churn, which focuses on a single relationship (lost vs retained customers) and works best in the application to contractual markets. This improvement included the four key behaviours within a public transport market: new riders, lost riders, retained riders, and returning riders as identified through the literature review. A brief description of the four segments based on ridership style is provided as follows:

- ***New riders:*** *Riders that are new to the transport mode within the given study period. Where possible, attempts will be made to measure first-time riders, but this is considered a small portion of the population (e.g. tourists).*

- ***Lost riders:*** *Those that have been consistently using a public transport mode and then stop using the mode and do not use it again within the period being studied*
- ***Retained riders****: Users that travel each month consistently with no more than 25% of the total study period as a break over non-consecutive months.*
- ***Returning riders:*** *Any user that rides sporadically in the period, consistently starting and stopping travel*.

The segments of new, lost and retained riders were based on the existing concept of customer churn as discussed by Riebe et al. (2014). In addition, the 'returning' category has been added to reflect the element of time and changing customer needs over time. This is based on the criticisms of churn posited by Tamaddoni et al. (2014) and Kamakura et al. (2005). These segments are all a function of user decisions to start, stop, or continue using a public transport mode.

The development of a framework that will measure the number of customers that start, continue, stop and return to travel using a public transport mode over time increases our existing understanding of net changes in public transport markets. This concept can also provide an understanding of the influences not just behind decisions to stop using a service, as is the current academic focus, but also reasons for starting, pausing, or continuing to travel. This has been recognized as a valuable contribution that may improve the accuracy of existing models (Ma et al., 2013, Chu, 2015).

To illustrate the value of measuring customer fluctuation within public transport markets, Figure 8.2 illustrates the way that new, lost, retained, and returning users interact within a market (as initially introduced in Chapter 3: Framework Development and Research Approach).

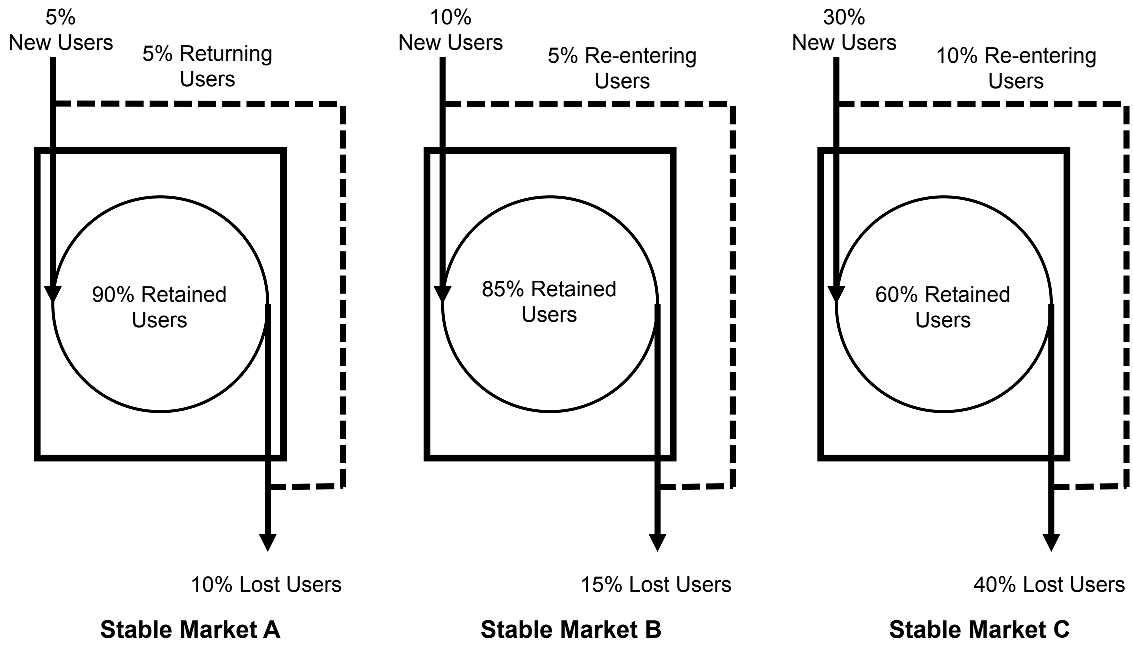

Figure 8.2 - Customer fluctuation within a Stable Market Context adapted by the author from Blythe (2009)

Based on current methods of measurement, all three of the markets pictured in Figure 3.4 would be considered the same, i.e. stable markets that are exhibiting no growth or decline. However, each of these markets is different, showing increasing rates of fluctuation with Market A having the least fluctuation and Market C having the highest levels of fluctuation.

A key question identified in the development of customer fluctuation is the appropriate temporal variables (measurement period and units of measurement) to capture changes in individual ridership. There was no consistency in the approaches taken by adjacent studies on public transport on the appropriate measurement period and time frames for measuring behaviour change. The measurement period varies from a single day to 5 years, and the unit of analysis varied from hours to months or even years in the studies reviewed in Chapter Two of this thesis (E.g. Briand et al., 2017, Chu, 2015, Morency et al., 2007, Ma et al., 2017, Tao et al., 2014). To appropriately measure customer fluctuation, the measurement period must be long enough to capture actual changes in ridership, rather than just seasonal variations, but not so long as to exaggerate change.

A one-year measurement period was proposed in the initial development of the concept, though it was found through testing in Chapter 5 that a two-year period, divided into 18 months of active analysis and a 3-month buffer at the beginning and end of the period, was the most appropriate approach. This is because it mitigates measurement period problems for new and lost users as well as seasonal variations in travel. The other important temporal variable is the unit of time used to measure change across the year; several approaches were tested in Chapter 4, with months determined as the most practical unit of analysis for this concept.

8.3.3 To explore a smart card and survey based approach to measure market change segments applied to the case of Melbourne (RO3)

This thesis developed an a-priori segmentation approach for defining customer fluctuation segments and then measured customer fluctuation using two approaches. One approach used smart card data, and the other was a primary survey of users, developed, applied and tested as part of the thesis.

8.3.3.1 Smart card approach to measuring market change

The smart card approach identified very high rates of customer fluctuation (a small proportion of retained cards and a high proportion of lost/returning cards) across all modes of public transport. This pattern of substantial fluctuation was generally consistent when using one year, two years, and 18 months (plus verification buffer) as the measurement period. Applying the measurement approach to one year of smart card data identified that it was difficult to differentiate the results from seasonal variations in public transport ridership. We concluded that customer fluctuation is best measured over a period longer than a year to ensure an accurate representation of changes in ridership.

The smart card approach provided a valuable opportunity to measure the total travel volume for each customer fluctuation segment. This found that though the retained segment was consistently the smallest proportion of smart cards across all modes, retained users were responsible for a much larger proportion of trips when compared to the proportion of users. Similarly, lost users were responsible for a lower proportion of total trips that when compared to the proportion of smart cards within the market. This analysis identified that even though markets might have a high share of lost users, this does not have a proportionate impact on the number of trips being taken. This finding highlights the importance of considering total travel volume when reviewing how customer fluctuation impacts the transport market; a key area for future research.

The smart card approach had a range of limitations. For example, it was not possible to determine the demographic characteristics of smart card users in different customer fluctuation segments. Technical limitations were also identified; for example, smart cards have an expiration date and therefore they may over-represent lost riders. Also, an unknown proportion of cards could be misplaced or shared between riders and some passengers use multiple smart cards. Fare evasion also obscures the link between card use and travel patterns.

8.3.3.2 Primary survey approach to measuring market change

The second approach was a primary survey. Estimates of customer fluctuation using this approach differed considerably from estimates found using the smart card approach. The survey approach found high levels of customer retention and low rates of customer defections (lost users) across all modes.

There were significant benefits associated with the primary survey approach. First, it made it possible to identify socio-demographic differences between the customer fluctuation segments. More significantly, it made it possible to explore behavioural factors influencing why people were in each market change segment. This is discussed further in Section 8.3.4 of this chapter.

There were also several limitations with the primary survey approach as surveys are expensive and complex to design well. Surveys were designed to be short and easy to complete, however as a result, participants only responded to questions for their ‘mainly used mode’ of public transport and not secondary modes where they might have patterns with greater fluctuation. This is likely to increase the impact of the existing response bias, where those who use public transport regularly are more likely to respond to a survey on the topic. Further, there was also a sampling bias in favour of public transport users, though there was a target of 25% of the sample for non-public transport users to account for lost users. The sample also failed to capture tourists or other temporary users that would be identified through smart card data.

The final limitation of survey data is that it relies on retrospective data, where participants were required to recall their travel behaviour over the last year. The analysis established evidence of a ‘social desirability bias’ impact survey results, with participants wanting to record themselves as public transport users despite actual travel patterns which do not involve transit use.

8.3.3.3 Comparison of Approaches

Overall, the two approaches shared some characteristics but were largely quite different. These differences are summarised in Figure 8.3.

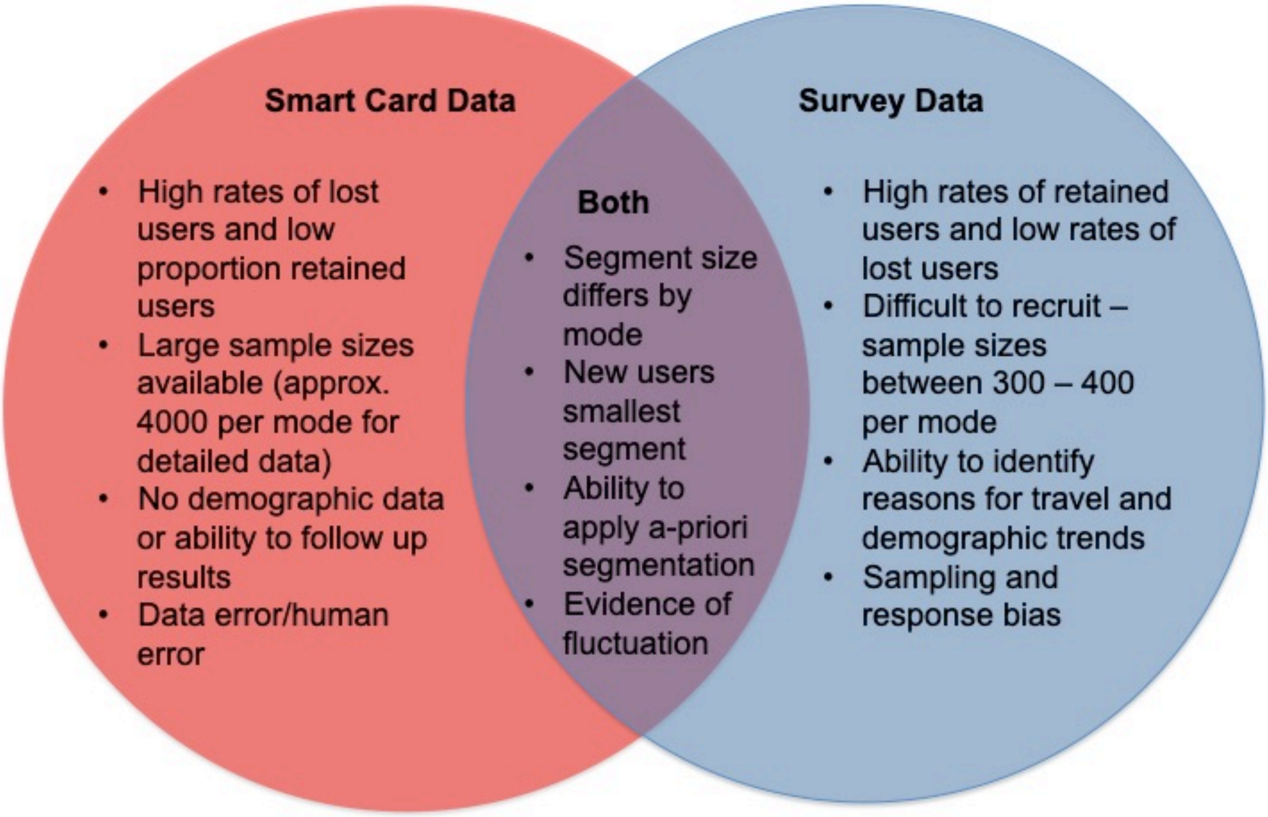

Figure 8.3 - Comparison of Smart Card and Survey Data Strengths and limitations

In both approaches variations by mode showed generally consistent patterns with train indicating the lowest levels of customer fluctuation (higher rates of retained/new/returning users and lower rates of lost users) and bus the highest rates of fluctuation (higher rates of lost users and lower rates of new/retained and returning users when compared to other modes). The analysis identified that the two year smart card data (with buffer) and verified survey results provided the most robust measurements of customer fluctuation. While results of the two methods were significantly different, each approach presented different strengths and limitations, it was identified that an integrated measurement approach is necessary.

8.3.4 To explore the behavioural factors influencing market change segments using survey data for metro Melbourne (RO4)

The primary survey approach provided the only avenue to investigate the reasons why individuals exhibit certain customer fluctuation behaviour. The range of potential factors that might influence behaviour was identified based on a review of the literature and included life events (Beige and Axhausen, 2012, Verplanken et al., 2008, Clark et al., 2014), lifestyle choices (Abrahamse et al., 2009, Bamberg et al., 2007, Jacques et al., 2013), the train/bus/tram service (van Lierop and El-Geneidy, 2016) and the trip taken (Bamberg et al., 2003, Buehler et al., 2017).

The research found that factors influencing customer fluctuation differed between customer fluctuation segments, and included the following:

- A complex mix of factors influenced new user behaviour. Personal factors like 'saving money' and 'don't have a car/ can't drive' were the most influential overall though life events like 'changed home/work locations' are also important. New use of rail was most influenced by trip taken factors such as 'parking difficulties' and 'travelling to an event'. New use of bus was influenced more by personal factors, notably 'don't own/drive a car' and the life event 'changed home/work location' is the most common new bus user influence.
- Lost user behaviour was predominantly influenced by personal factors like 'I no longer needed to make trips with the train/bus/tram'. This was followed by life events, which was driven by those that selected 'I went on holiday' and 'my life circumstances changed'. Service factors were the third most selected, led by a proportion that selected 'catching the train/bus/tram takes too long’. There is evidence that if a more convenient option becomes available, people will defect from public transport use. The results also suggest that service factors were likely to have a limited impact on stopping customer defection.
- Returning users that paused travel showed similar motivations to lost users and were overall driven by personal factors. Again, the core motivation for pausing travel were around no longer needing to use public transport. However, the reasons varied significantly for bus users compared to other modes, with service factors the primary reason influencing travel decisions.
- Among returning users that recommenced travel, the trip taken was the leading group of reasons for all modes. All modes had a significant proportion of responses around 'I used the train/bus/tram when travelling to an event', 'parking is too difficult at my destination', and 'the train/bus/tram was the most convenient option'. These responses indicate that special event (e.g. large sporting or entertainment events) or location-based travel (travel to the city) was a key reason returning users recommence travel.
- Retained users for train and tram use are primarily influenced by service factors, whereas retained bus users are mainly influenced by personal factors e.g. not owning a car or being unable to drive. Retained users for all modes are primarily influenced by the convenience of public transport options.

It is evident in the summary of behavioural factors above that external factors, particularly those related to the private car, were the most influential. Service factors, which public transport operators can influence, had less impact on ridership.

8.3.5 To assess the potential of the new approach for application in the industry (RO5)

The research approach and analysis presented in this thesis found many difficulties with the practicalities of designing and implementing the measurement of customer fluctuation segments. It is likely that these difficulties explain why a gap in knowledge exists in this area.

The key practical difficulties encountered included:

- The need for a substantial measurement period, greater than a single year, to ensure customer fluctuation is being captured rather than seasonal variations in travel.
- Limited comparable studies available which could be used to assist in verifying the customer fluctuation results of this approach.
- In Victoria, there were no existing measurement tools (for example, the Tracker survey discussed in Chapter 3) that could be used to measure customer fluctuation without significant adjustments.
- The market change segments identified were all sensitive to the temporal unit of measurement selected and there was no consensus on appropriate temporal units in the adjacent public transport literature.
- The data sources available – both primary and secondary - presented significant limitations for the measurement of customer fluctuation. Smart card data presented inherent data collection errors as well as user errors, which were identified as likely to exaggerate customer fluctuation. Primary data included response and social desirability biases that were likely to make the measurement of customer fluctuation segments imprecise.

Though this analysis has identified several limitations with this the first application of customer fluctuation to public transport markets, the findings still suggest that this approach has significant potential, both observed and anticipated. The potential of this approach is briefly outlined below:

- Early results have identified that there are significant levels of customer fluctuation occurring within Melbourne’s public transport markets. These findings, though subject to the limitations discussed, could not be replicated with any existing measurement tools within the case of Melbourne and warrant further refinement and investigation.

- Reviewing customer fluctuation segments in terms of total travel volume, rather than individual riders or cards, would provide additional significant insights into just how influential different market segments are in relation to ridership growth.
- Insights from regularly measuring customer fluctuation, through either smart card or survey data, as a tool to track changes in market change segment size over time. This would allow operators to identify their benchmark level of customer fluctuation and be aware of any changes to segment size that might signal a potential issue or opportunity.
- There is potential to apply customer fluctuation to different public transport markets for comparison. This might be a valuable way to create and understand industry benchmarks for the level of fluctuation that is occurring within a market. Further, this information may be used to help build a case for new public transport projects.
- The approach offers the potential to deepen understanding of why public transport users are choosing to change travel behaviour. The application of customer fluctuation using a survey identified that motivations influencing travel behaviour were primarily exogenous to public transport operations for all ridership segments. This is apart from the retained segment, as users that have already made the decision to use public transport where more likely to be impacted to internal and service changes. The ability to complete further study and discern direct links between segment characteristics and marketing strategies or service modification is likely to be of great relevance for public transport operators.

Despite the practical limitations, the potential of this approach warrants the investigation of combined measurement tools. Chapter 7 suggested a refined approach that combines both smart card data and survey data to create a more complete method of measurement. This could be through an existing panel of smart card users run by an organisation, or through an independent study. The use of a panel would ensure the ability to track longitudinal data from specific individuals' smart cards while also being able to verify and understand travel through participant interviews or surveys. Similar approaches have been utilised by Ji et al. (2019), Medina and Arturo (2018), Li et al. (2018b).

8.4 Research Contributions

This research developed a new concept for the measurement of market change analysis and applied it to Melbourne's public transport markets. As such, there were several new contributions to the field of public transport and marketing research that have been made through this thesis. These contributions have been summarised in Table 8.2 and are shown in relation to relevant research objectives and chapter locations where these contributions are made within this thesis.

Table 8.2 - Summary of Research Contributions relevant to each Research Objective

Research Objectives	***Contributions***
RO1: To understand conventional measures of market change and explore the benefits and drawbacks of these methods (Chapter 2)	• The application of the four elements of churn to public transport markets as a new approach for measuring market change through a single measurement approach
RO2: To develop a new concept for market change analysis based on market change segments (Chapter 3)	• The development of a mixed method approach to test the measurement of customer fluctuation – a new method for measuring market change in public transport • The development of a-priori segmentation rules that identify the temporal definitions for each fluctuation segment and allow for comparison between measurement approaches
RO3: To explore a smart card and survey based approach to measure market change segments applied to the case of metro Melbourne (Chapter 4, 5 and 6)	• The research identified inherent limitations with smart card data that cannot be addressed through this method and limit its potential as an independent data source for measuring market change • There are significant variations in change segment measurement between measurement approaches using smart card data and primary survey data. The research identified a need to develop a more integrated approach capturing the benefits of both alternatives.
RO4: To explore behavioural factors influencing market change segments using survey data for metro Melbourne (Chapter 6)	• The research suggested that the common reasons influencing fluctuation behaviour were external to public transport markets and outside of the control of operators. However, there were service factors that could be improved to keep retained users and public transport operators might target advocacy to reduce car dependency such as reduced parking coupled with increased services to events.
RO5: To assess the potential of the new approach for application in the industry (Chapter 7)	• The research has found that ongoing fluctuations in market change segments are difficult to measure in practice and may be too complex and costly to implement without further research and testing of new approaches. • The research recommends the development of a new integrated mixed methods approach to reduce the data and technical limitations of both smart card and survey data measurement methods.

These contributions represent an important contribution to knowledge through the findings and research completed as part of this thesis.

8.5 Research Limitations

This thesis presents the development and exploration of a new framework – customer fluctuation. The research methods employed in this initial exploration of customer fluctuation produced results that met the research aims and objectives set for this thesis. Despite this, there are several key research limitations that must be noted within the research method presented.

The most significant limitation for this research was related to the data sources that were utilised in the analysis. Both the smart card and primary survey approaches adopted had significant limitations impacting on the measurement results.

For smart card data, the key limitations were identified as follows:

- Smart cards cannot be equated to individual riders because users may hold multiple smart cards or lose and need new smart cards regularly.
- ‘myki’ smart cards expire after four years of purchase, requiring a new card with a new ID. These IDs are not linked between new and old cards. These expiry rates might also vary when looking at different concession requirements.
- Smart card data cannot perfectly capture the true number of riders due to the prevalence of fare evasion and shared cards.
- Smart card data is anonymised and there is difficulty linking this data to demographics.

These limitations are inherent within smart card data and for most public transport agencies using similar approaches. These limitations cannot be mitigated despite the large data sample utilised.

The use of smart card data was also limited in its scope, as the research methods did not include co-ordinate data of trips (e.g. touch on and touch off locations) which might have been analysed for smart card based insights into travel behaviour. This would require a complex trip-chaining process and was not considered within this project due to time constraints but would be a valuable area for future research.

There were also several limitations with the primary survey design and data that must be considered. The survey design provided descriptive statements relating to the customer fluctuation segments for participants to self-categorise their travel patterns into market change segments. Participants struggled to self-categorise as their chosen description did not always match their stated behaviour. A solution was identified using ‘verified’ travel patterns of survey respondents, suggesting this is a more robust approach to segment classification for future research.

The primary survey approach also had limitations associated with the financial and time constraints of the project. A sample size to allow for 95% accuracy (+/ - 5) was not universally achieved for all modes. Further, due to low responses for tram and bus users, the data for these modes was collected over a longer period than for train users. This may have affected the comparability of results between modes. In addition, participants were only asked to respond for their ‘mainly used mode’ which may act to overemphasise retained users and does not account for the impact of multi-modal travel on the customer fluctuation rates of individual modes. Finally, the survey used retrospective data, which relied on accurate participant recall of travel behaviour over the previous year. As the analysis of customer fluctuation found that a longer measurement period is required, this calls into question the validity of this approach as it is unlikely participants could accurately recall their behaviour perfectly over a longer period. This may be modified with further questions around their confidence in their responses.

The customer fluctuation segmentation approach explored in this thesis is the first application of this concept and requires further refinement. Both the smart card and primary survey approaches found that the segments identified were very sensitive to changes in the temporal unit of analysis as well as the measurement period.

All market change segments require further investigation and refinement to ensure they can be applied in multiple contexts. The returning user segment is currently categorised as any user that does not fit within the categories of new, lost, or retained. Future applications of customer fluctuation may wish to further refine this category and potentially introduce an additional segment to give more insight into this broad category.

Overall, this thesis did not find a single conclusive method for measuring customer fluctuation. Rather, advances have been made in identifying possible approaches and in finding limitations in those approaches. This will make future development of an improved approach more feasible. The research did establish that customer fluctuation is a significant issue and warrants more investigation to make robust measurement approaches available into the future. The smart card approach found evidence that over half of public transport users for all modes are lost over a two-year period, and that there is only a small share of retained users. Conversely, the primary survey found low rates of lost users and high rates of retention within Melbourne markets. As it is likely that the correct results sit somewhere between these two measurements, it is evident that customer fluctuation is a complex process that is occurring within public transport markets

8.6 Areas for Further Research

This thesis represents the first time the concept of customer fluctuation is developed, applied, and measured. As might be expected with anything new, many practical difficulties and limitations were found. It is suggested that further research continues with an exploration of more robust measurement approaches and builds upon and refines the ideas presented in this thesis.

Modelling to confirm the rate of customer fluctuation in transport markets and test the accuracy of the a-priori segmentation approach is a recommended area for future research. This should include both statistical transport models such as hazard modelling (Trépanier et al., 2012) or discrete choice models (Bass et al., 2011) and clustering techniques such as DBSCAN plus noise (Ma et al., 2013). It is recommended that modelling is completed using the hybrid data approach developed in Chapter 7. It is worth re-iterating that a focus of this research is developing a model or tool for understanding patterns of customer fluctuation that is meaningful for practitioners, a gap noted by Yoh et al, (2012). As such, modelling should be used to test and refine, though not replace, the customer fluctuation approach identified in this thesis.

It is noted that despite potential accuracy issues, if public transport operators were to regularly measure customer fluctuation, through either smart card or survey data, it could be a valuable tool to track changes in market segment size over time. This would allow operators to identify their benchmark level of customer fluctuation and be aware of any changes to segment size. This might be used to flag potential issues and allow operators to ask targeted questions through surveys or existing panels.

Refinement of the customer fluctuation measurement approach should also establish total travel volume as a key measure rather than the number of people who travel. This is likely to hold much more significance for public transport operators as it is more directly related to market size.

Finally, like studies of customer churn in marketing, further analysis to develop models which help predict customer fluctuation should be completed (Bellman et al., 2010, Holtrop et al., 2016, Neslin et al., 2006, Tsai and Lu, 2009, Zhao et al., 2005). Prediction of travel behaviour is also a focus on existing public transport research (Oort et al., 2015, Zhao et al., 2018). A predictive model would allow for the ongoing monitoring of public transport markets as well as the ability to test marketing interventions with segmentation results compared to predictive data. As an example, we found evidence that there was a direct relationship between car ownership and lost bus users and a relationship between large scale (sporting) events and returning users for all modes. Questions might be included in regular travel surveys to help predict the impact of these or similar factors.

8.7 COVID-19 Impacts Codicil

The data collection and analysis for this thesis has all been conducted prior to the impacts of COVID-19 being felt in Australia (March 2020). As such, this research has not addressed the unprecedented impacts of a global public health event on the public transport industry. It is considered that in the face of greatly reduced public transport use and capacity as well as changing work and overall patterns that customer fluctuation has great potential as a concept to measure the internal changes within public transport markets. Where local public transport settings have sufficient data, likely smart card, to measure previously occurring customer fluctuation this would allow for a unique measurement of the changes occurring within public transport markets both during and recovering from the COVID-19 pandemic. This is suggested as an area for future research with some of the most significant potential to provide new insights for public transport operators.

8.8 Concluding Remarks

This thesis explored the new concept of customer fluctuation as a new means of understanding market change in public transport. Growing public transport use is a fundamental goal of public transport operators globally. Despite this, there are few studies that seek to understand what changes are happening to individual ridership within the public transport market. This research found evidence that there are moderate to high levels of market change occurring within Melbourne’s public transport markets, and that this is impacting all modes. This research also identified that the reasons behind these variations are only partly influenced by service attributes or other factors within public transport operators' control. It is suggested that further refining and incorporating customer fluctuation measurements into the analysis of public transport studies may have significant benefits for helping to understand and grow markets into the future.

References

Abou-Zeid, M. & Ben-Akiva, M. 2012. Travel mode switching: Comparison of findings from two public transportation experiments. *Transport Policy,* 24, 48-59.

Abou-Zeid, M. & Fujii, S. 2016. Travel satisfaction effects of changes in public transport usage. *Transportation,* 43, 301-314.

Abou-Zeid, M., Witter, R., Bierlaire, M., Kaufmann, V. & Ben-Akiva, M. 2012. Happiness and travel mode switching: Findings from a Swiss public transportation experiment. *Transport Policy,* 19, 93-104.

Abrahamse, W., Steg, L., Gifford, R. & Vlek, C. 2009. Factors influencing car use for commuting and the intention to reduce it: A question of self-interest or morality? *Transportation Research Part F: Traffic Psychology and Behaviour,* 12, 317-324.

Ahmad Termida, N., Susilo, Y. O., Franklin, J. P. & Liu, C. 2018. Understanding seasonal variation in individual's activity participation and trip generation by using four consecutive two-week travel diary. *Travel Behaviour and Society,* 12, 52-63.

Ahn, J.-H., Han, S.-P. & Lee, Y.-S. 2006. Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications policy,* 30, 552 - 568.

Allen, D. R. 2004. *Customer satisfaction research management : a comprehensive guide to integrating customer loyalty and satisfaction metrics in the management of complex organizations,* Milwaukee, ASQ Quality Press

Alsger, A., Tavassoli, A., Mesbah, M., Ferreira, L. & Hickman, M. 2018. Public transport trip purpose inference using smart card fare data. *Transportation Research Part C: Emerging Technologies,* 87, 123-137.

Amiripour, S. M. M., Ceder, A. & Mohaymany, A. S. 2014. Designing large-scale bus network with seasonal variations of demand. *Transportation Research Part C: Emerging Technologies,* 48, 322-338.

Anable, J. 2005. 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. *Transport policy,* 12, 65-78.

Ascarza, E. & Hardie, B. G. S. 2013. A Joint Model of Usage and Churn in Contractual Settings. *Marketing Science,* 32, 570-590.

Athanassopoulos, A. D. 2000. Customer Satisfaction Cues To Support Market Segmentation and Explain Switching Behavior. *Journal of Business Research,* 47, 191-207.

Australian Bureau of Statistics. 2018a. *2016 Census Community Profiles* [Online]. Canberra: Australian Bureau of Statistics. Available: https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/communityprofile/2GMEL?opendocument [Accessed 8 April 2019

Australian Bureau of Statistics. 2018b. *3218.0 - Regional Population Growth, Australia, 2016-17* [Online]. Canberra: Australian Bureau of Statistics. Available: https://www.abs.gov.au/ausstats/abs@.nsf/mf/3218.0 [Accessed 25 Feburary 2019].

Australian Bureau of Statistics. 2018c. *What's driving population growth in Australia's cities?* [Online]. Canberra: Australian Bureau of Statistics. Available: http://www.abs.gov.au/ausstats/abs@.nsf/latestProducts/3218.0Media Release12016-17?OpenDocument [Accessed 15 August 2018].

Axhausen, K., Löchl, M., Schlich, R., Buhl, T. & Widmer, P. 2007. Fatigue in long-duration travel diaries. *Transportation,* 34, 143-160.

Bagchi, M. & White, P. R. 2005. The potential of public transport smart card data. *Transport Policy,* 12**,** 464-474.

Ballings, M. & Van den Poel, D. 2012. Customer event history for churn prediction: How long is long enough? *Expert Systems with Applications,* 39**,** 13517-13522.

Bamberg, S., Ajzen, I. & Schmidt, P. 2003. Choice of Travel Mode in the Theory of Planned Behavior: The Roles of Past Behavior, Habit, and Reasoned Action. *Basic and Applied Social Psychology,* 25**,** 175-187.

Bamberg, S., Hunecke, M. & Blöbaum, A. 2007. Social context, personal norms and the use of public transportation: Two field studies. *Journal of Environmental Psychology,* 27**,** 190-203.

Bass, P., Donoso, P. & Munizaga, M. 2011. A model to assess public transport demand stability. *Transportation Research Part A: Policy and Practice,* 45**,** 755-764.

Behrens, R. & Mistro, R. D. 2010. Shocking Habits: Methodological Issues in Analyzing Changing Personal Travel Behavior Over Time. *International Journal of Sustainable Transportation,* 4**,** 253-271.

Beige, S. & Axhausen, K. 2012. Interdependencies between turning points in life and long-term mobility decisions. *Planning - Policy - Research - Practice,* 39**,** 857-872.

Beige, S. & Axhausen, K. W. 2008. Long-term and Mid-term Mobility Decisions During the Life Course: Experiences with a Retrospective Survey. *IATSS Research,* 32**,** 16-33.

Beimborn, E., Greenwald, M. & Jin, X. 2003. Accessibility, connectivity, and captivity: impacts on transit choice. *Transportation Research Record: Journal of the Transportation Research Board***,** 1-9.

Bellman, S., Jarvis, W. & Zorn, S. 2010. Attitudinal perspectives for predicting churn. *Journal of Research in Interactive Marketing,* 4**,** 157-169.

Bhat, C. & Srinivasan, S. 2005. *A Multidimensional Mixed Ordered-Response Model for Analyzing Weekend Activity Participation*.

Blythe, J. 2009. *Key Concepts in Marketing,* London, UNKNOWN, SAGE Publications.

Bocker, L., Dijst, M. & Prillwitz, J. 2013. Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review. *Transport Reviews,* 33**,** 71 - 91.

Bolton, R. N. 1998. A Dynamic Model of the Duration of the Customer's Relationship With a Continuous Service Provider: The Role of Satisfaction. *Marketing Science,* 17**,** 45.

Bolton, R. N., Kannan, R. K. & Bramlett, M. D. 2000. Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value. *Journal of the Academy of Marketing Science,* 28**,** 95.

Brady, D. 2014. *A Better Way to Calculate Your Churn Rate* [Online]. Recurly. Available: https://blog.recurly.com/better-way-to-calculate-your-churn-rate [Accessed 20 April 2019].

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

Briand, A. S., Côme, E., Mahrsi, M. K. E. & Oukhellou, L. A mixture model clustering approach for temporal passenger pattern characterization in public transport. 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 19-21 Oct. 2015 2015. 1-10.

Buckinx, W. & Van den Poel, D. 2005. Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research,* 164**,** 252-268.

Buehler, R., Pucher, J., Gerike, R. & Götschi, T. 2017. Reducing car dependence in the heart of Europe: lessons from Germany, Austria, and Switzerland. *Transport Reviews,* 37**,** 4-28.

Chalak, A., Al-Naghi, H., Irani, A. & Abou-Zeid, M. 2016. Commuters' behavior towards upgraded bus services in Greater Beirut: Implications for greenhouse gas emissions, social welfare and transport policy. *Transportation Research Part A: Policy and Practice,* 88**,** 265-285.

Chatterjee, K. 2011. Modelling the dynamics of bus use in a changing travel environment using panel data. *Transportation,* 38**,** 487-509.

Chen, C., Varley, D. & Chen, J. 2011a. What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour. *Urban Studies,* 48**,** 1893-1908.

Chen, C., Varley, D. & Chen, J. 2011b. What Affects Transit Ridership? A Dynamic Analysis involving Multiple Factors, Lags and Asymmetric Behaviour *Urban Studies,* 48**,** 1893-1908.

Chu, K. K. A. 2015. Two-year Worth of Smart Card Transaction Data – Extracting Longitudinal Observations for the Understanding of Travel Behaviour. *Transportation Research Procedia,* 11**,** 365-380.

City of Melbourne. *Melbourne facts and figures* [Online]. Melbourne: City of Melbourne. Available: https://www.melbourne.vic.gov.au/about-melbourne/melbourne-profile/Pages/facts-about-melbourne.aspx [Accessed 15 August 2018].

Clark, B., Chatterjee, K., Melia, S., Knies, G. & Laurie, H. 2014. Life Events and Travel Behavior: Exploring the Interrelationship Using UK Household Longitudinal Study Data. *Transportation Research Record,* 2**,** 54-64.

Clayton, W., Jain, J. & Parkhurst, G. 2016. An ideal journey: making bus travel desirable. *Mobilities***,** 1-20.

Connidis, I. 1983. Integrating qualitative and quantitative methods in survey research on aging: An assessment. *Qualitative Sociology,* 6**,** 334-352.

Cooper, J. & Corcoran, J. 2018. *2071.0.55.001 - Census of Population and Housing: Commuting to Work - More Stories from the Census, 2016, 2016* [Online]. Canberra: Australian Bureau of Statistics,. Available: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/bySubject/2071.0.55.001~2016~Main Features~Feature Article: Journey to Work in Australia~40 [Accessed 27 October 2019.

Csikos, D. & Currie, G. 2008. Investigating Consistency in Transit Passenger Arrivals: Insights from Longitudinal Automated Fare Collection *Transportation Research Record: Journal of the Transportation Research Board,* No. 2042**,** 12 - 19.

Currie, G. & Wallis, I. 2008. Effective ways to grow urban bus markets - a synthesis of evidence. *Journal of Transport Geography,* 16**,** 419 - 430.

Davidov, E. 2007. Explaining Habits in a New Context the Case of Travel-Mode Choice. *Rationality and society,* 19**,** 315-334.

De Vos, J., Mokhtarian, P. L., Schwanen, T., Van Acker, V. & Witlox, F. 2016. Travel mode choice and travel satisfaction: bridging the gap between decision utility and experienced utility. *Transportation,* 43**,** 771-796.

Dempsey, P. S. 2015. *Privacy Issues with the Use of Smart Cards* [Online]. Available: SSRN: https://ssrn.com/abstract=2692652 or http://dx.doi.org/10.2139/ssrn.2692652.

Denzin, N. K. 1978. The Research Act. 2nd ed. New York: McGraw-Hill.

Devillaine, F., Munizaga, M. & Trépanier, M. 2012. Detection of Activities of Public Transport Users by Analyzing Smart Card Data. *Transportation Research Record,* 2276**,** 48-55.

Dick, A. & Basu, K. 1994a. Customer loyalty: Toward an integrated conceptual framework. *Journal of the Academy of Marketing Science***,** 99.

Dick, A. S. & Basu, K. 1994b. Customer Loyalty: Toward an Integrated Conceptual Framework. *Journal of the Academy of Marketing Science,* 22**,** 99-113.

Dowling, G. R. & Uncles, M. 1997. Do Customer Loyalty Programs Really Work? *Sloan Management Review,* 38**,** 71-82.

Duy, Q., Currie, G., Gruyter, C. & Young, W. 2018. Exploring the impact of public transport strikes on travel behavior and traffic congestion. *International Journal of Sustainable Transportation***,** 1-11.

Ehrenberg, A. S. C. 1972. *Repeat-buying; theory and applications,* New York, American Elsevier.

Elmore-Yalch, R. 1998. *A handbook: using market segmentation to increase transit ridership,* Washington, D.C. , National Academy Press

Glady, N., Baesens, B. & Croux, C. 2009. Modeling churn using customer lifetime value. *European Journal of Operational Research,* 197**,** 402-411.

Goulet Langlois, G., Koutsopoulos, H. N. & Zhao, J. 2016. Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C: Emerging Technologies,* 64**,** 1-16.

Guiver, J. W. 2007. Modal talk: Discourse analysis of how people talk about bus and car travel. *Transportation Research Part A: Policy and Practice,* 41**,** 233-248.

Hoang-Tung, N., Kojima, A. & Kubota, H. 2014. Recognising customers' patterns of bus service patronage using a loyalty framework. *IATSS Research,* 37**,** 148-156.

Holtrop, N., Wieringa, J. E., Gijsenberg, M. J. & Verhoef, P. C. 2016. No future without the past? Predicting churn in the face of customer privacy. *International Journal of Research in Marketing*.

Jacques, C., Manaugh, K. & El-geneidy, A. M. 2013. Rescuing the captive [mode] user: an alternative approach to transport market segmentation. *Transportation,* 40**,** 625-645.

Ji, Y., Cao, Y., Liu, Y., Guo, W. & Gao, L. 2019. Research on classification and influencing factors of metro commuting patterns by combining smart card data and household travel survey data. *IET Intelligent Transport Systems,* 13**,** 1525-1532.

Jick, T. D. 1979. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly,* 24**,** 602-611.

Jin, X., Beimborn, E. & Greenwald, M. 2005. Impacts of accessibility, connectivity and mode captivity on transit choice. Milwaukee, WI: Center for Urban Transportation Studies: University of Wisconsin-Milwaukee.

Kamakura, W., Mela, C., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., Naik, P., Neslin, S., Sun, B., Verhoef, P., Wedel, M. & Wilcox, R. 2005. Choice Models and Customer Relationship Management. *Marketing Letters,* 16**,** 279 - 291.

Kieu, L.-M., Bhaskar, A. & Chung, E. 2015a. A modified Density-Based Scanning Algorithm with Noise for spatial travel pattern analysis from Smart Card AFC data. *Transportation Research Part C: Emerging Technologies,* 58**,** 193-207.

Kieu, L. M., Bhaskar, A. & Chung, E. 2014. Transit passenger segmentation using travel regularity mined from Smart Card transactions data. *Transportation Research Board 93rd Annual Meeting.* Washington, D.C.

Kieu, L. M., Bhaskar, A. & Chung, E. 2015b. Passenger Segmentation Using Smart Card Data. *IEEE Transactions on Intelligent Transportation Systems,* 16**,** 1537-1548.

Kim, J., Corcoran, J. & Papamanolis, M. 2017. Route choice stickiness of public transport passengers: Measuring habitual bus ridership behaviour using smart card data. *Transportation Research Part C: Emerging Technologies,* 83**,** 146-164.

Kitamura, R. & Hoorn, T. 1987. Regularity and irreversibility of weekly travel behavior. *Planning - Policy - Research - Practice,* 14**,** 227-251.

Kotler, P. 1994. *Marketing Management,* Upper Saddle River, NJ, Prentice Hall.

Krizek, K. & El-Geneidy, A. 2007. Segmenting preferences and habits of transit users and non-users. *Journal of Public Transport,* 10**,** 71-94.

Lai, W.-T. & Chen, C.-F. 2011. Behavioral intentions of public transit passengers - The roles of service quality, percieved value, satisfaction and involvement. *Transport Policy,* 18**,** 318 - 325.

Lee, J., Lee, J. & Feick, L. 2001. The impact of switching costs on the customer satisfaction-loyalty link: mobile phone service in France. *The Journal of Services Marketing,* 15**,** 35-48.

Li, T., Sun, D., Peng, J. & Yang, K. 2018a. Smart Card Data Mining of Public Transport Destination: A Literature Review. *Information,* 9**,** 18.

Li, Y.-T., Iwamoto, T., Schmöcker, J.-D., Nakamura, T. & Uno, N. 2018b. Analyzing long-term travel behaviour: A comparison of smart card data and graphical usage patterns. *Transportation Research Procedia,* 32**,** 34-43.

Libai, B., Muller, E. & Peres, R. 2009. The Diffusion of Services. *Journal of Marketing Research,* 46**,** 163-175.

Ľudmila Jánošíková, J. S. M. K. 2014. Estimation of a route choice model for urban public transport using smart card data. *Transportation Planning and Technology,* 37**,** 638-648.

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

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

Mason, A., Segal, J. & Condry, B. 2011. Market Churn in the British Rail Passenger Commuter and Leisure Markets. *European Transport Conference.* Glasgow, United Kingdom.

Matthies, E., Klöckner, C. A. & Preißner, C. L. 2006. Applying a Modified Moral Decision Making Model to Change Habitual Car Use: How Can Commitment be Effective? *Applied Psychology: An International Review,* 55**,** 91-106.

Medina, O. & Arturo, S. 2018. Inferring weekly primary activity patterns using public transport smart card data and a household travel survey. *Travel Behaviour and Society,* 12**,** 93-101.

Merkert, R. & Webber, T. 2018. How to manage seasonality in service industries – The case of price and seat factor management in airlines. *Journal of air transport management,* 72**,** 39-46.

Morency, C., Trepanier, M. & Agard, B. 2006. Analysing the variability of transit users behaviour with smart card data. *2006 IEEE Intelligent Transportation Systems Conference.* Toronto, Canada

Morency, C., Trépanier, M. & Agard, B. 2007. Measuring transit use variability with smart-card data. *Transport Policy,* 14**,** 193-203.

Morgan, R. 1992. Does the leaky bucket theory hold water? (Adweek study indicates that instability of advertising business is vastly overestimated). *ADWEEK Eastern Edition,* 33**,** 46.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J. & Mason, C. H. 2006. Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models. *Journal of Marketing Research,* XLIII**,** 204 - 211.

Oliver, R. L. 1999. Whence Consumer Loyalty? *Journal of Marketing,* 63**,** 33-44.

Olsson, L. E., Gärling, T., Ettema, D., Friman, M. & Fujii, S. 2013. Happiness and Satisfaction with Work Commute. *Social Indicators Research,* 111**,** 255-263.

Oort, N. V., Brands, T. & de Romph, E. 2015. Short-Term Prediction of Ridership on Public Transport with Smart Card Data. *Transportation Research Record: Journal of the Transportation Research Board,* 2535**,** 105-111.

Ortega-Tong, M. A. 2013. *Classification of London's Public Transport Users Using Smart Card Data.* Master of Science in Transportation, Massachusetts Institute of Technology

Pelletier, M.-P., Trépanier, M. & Morency, C. 2011. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies,* 19**,** 557-568.

Public Transport Victoria 2017. Annual Report 2016–17 Victoria: Public Transport Victoria.

Public Transport Victoria 2018. Victorian Fares and Ticketing Manual. Melbourne: Finsbury Green.

Reichheld, F. F. 1996. *The Loyalty Effect: The Hidden Force behind Growth, Profits and Lasting Value,* Boston, MA, Harvard Business School Press.

Reichheld, F. F. & Sasser, W. E. 1990. *Zero Defections: Quality Comes to Services* [Online]. Boston, MA: Harvard Business School Publishing. Available: https://hbr.org/1990/09/zero-defections-quality-comes-to-services [Accessed 4 September 2016.

Reinartz, W., Thomas, J. S. & Kumar, V. 2005. Balancing Acquisition and Retention Resources to Maximize Customer Profitability. *Journal of Marketing,* 69**,** 63-79.

Reinartz, W. J. & Kumar, V. 2000. On the profitability of long-life customers in a noncontractual Setting: An empirical investigation and implementations for marketing. *Journal of Marketing,* 64**,** 17 - 35.

Reinartz, W. J. & Kumar, V. 2003. The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing,* 67**,** 77-99.

Riebe, E., Wright, M., Stern, P. & Sharp, B. 2014. How to grow a brand: Retain or acquire customers? *Journal of Business Research,* 67**,** 990-997.

Rindfleisch, A., Malter, A. J., Ganesan, S. & Moorman, C. 2008. Cross-Sectional versus Longitudinal Survey Research: Concepts, Findings, and Guidelines. *Journal of Marketing Research,* 45**,** 261-279.

Saleh, W. & Farrell, S. 2007. Investigation and analysis of evidence of asymmetric churn in travel demand models. *Transportation Research Part A: Policy and Practice,* 41**,** 691-702.

Schlich, R. & Axhausen, K. W. 2003. Habitual travel behaviour: Evidence from a six-week travel diary. *Transportation,* 30**,** 13-36.

Schmitt, L., Currie, G. & Delbosc, A. 2013. Measuring the impact of unfamiliar transit travel using a university access survey. *Transport Policy,* 30**,** 301-307.

Schmitt, L., Currie, G. & Delbosc, A. 2015. Lost in transit? Unfamiliar public transport travel explored using a journey planner web survey. *Transportation,* 42**,** 101-122.

Schneider, B. & Bowen, D. 1985. Employee and Customer Perceptions of Service in Banks: Replication and Extension. *Journal of Applied Psychology,* 70**,** 423 - 433.

Sharp, B., Wright, M. & Goodhardt, G. 2002. Purchase Loyalty is Polarised into either Repertoire or Subscription Patterns. *Australasian Marketing Journal (AMJ),* 10**,** 7-20.

St-Louis, E., Manaugh, K., van Lierop, D. & El-Geneidy, A. 2014. The happy commuter: A comparison of commuter satisfaction across modes. *Transportation Research Part F: Traffic Psychology and Behaviour,* 26, Part A**,** 160-170.

State of Victoria 2016 Victoria in Future 2016: Population and household projections to 2051. *In:* Department of Environment, L., Water and Planning (ed.). Melbourne: State of Victoria Department of Environment, Land, Water and Planning

Survey Monkey. 2019. *Sample size calculator* [Online]. Survey Monkey. Available: https://www.surveymonkey.com/mp/sample-size-calculator/ [Accessed 15 November 2019].

Susilo, Y. & Axhausen, K. 2014. Repetitions in individual daily activity–travel–location patterns: a study using the Herfindahl–Hirschman Index. *Planning - Policy - Research - Practice,* 41**,** 995-1011.

Taherdoost, H. 2017. Determining Sample Size; How to Calculate Survey Sample Size.

Tamaddoni, A., Stakhovych, S. & Ewing, M. 2016. Comparing Churn Prediction Techniques and Assessing Their Performance: A Contingent Perspective *Journal of Service Research* 19, 123 - 141.

Tamaddoni, A. J., Sepehri, M. M., Teimourpour, B. & Choobdar, S. 2010. Modeling customer churn in a non-contractual setting: the case of telecommunications service providers. *Journal of Strategic Marketing,* 18, 587-598.

Tamaddoni, A. J., Stakhovych, S. & Ewing, M. 2014. Managing B2B customer churn, retention and profitability. *Industrial Marketing Management,* 43, 1258-1268.

Tang, L. & Thakuriah, P. 2012. Ridership effects of real-time bus information system: A case study in the City of Chicago. *Transportation Research Part C: Emerging Technologies,* 22, 146-161.

Tao, S., Corcoran, J. & Mateo-Babiano, I. 2017a. Modelling loyalty and behavioural change intentions of busway passengers: A case study of Brisbane, Australia. *IATSS Research*.

Tao, S., Corcoran, J. & Mateo-Babiano, I. 2017b. Modelling loyalty and behavioural change intentions of busway passengers: A case study of Brisbane, Australia. *IATSS Research,* 41, 113-122.

Tao, S., Rohde, D. & Corcoran, J. 2014. Examining the spatial–temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. *Journal of Transport Geography,* 41, 21-36.

Taylor, B. D. & Fink, C. N. Y. 2013. Explaining transit ridership: What has the evidence shown? *Transportation Letters,* 5, 15-26.

Taylor, M. A. P. 2007. Voluntary Travel Behavior Change Programs in Australia: The Carrot Rather Than the Stick in Travel Demand Management. *International Journal of Sustainable Transportation,* 1, 173-192.

Thøgersen, J. 2006. Understanding repetitive travel mode choices in a stable contect: A panel study approach. *Transportation Research Part A,* 40, 621 - 638.

Transit Cooperative Research Program 1999. TCRP Report 47: A Handbook for Measuring Customer Satisfaction and Service Quality. Washington: Transportation Research Board.

Trépanier, M., Habib, K. M. N. & Morency, C. 2012. Are transit users loyal? Revelations from a hazard model based on smart card data. *Canadian Journal of Civil Engineering,* 39, 610-618.

Tsai, C.-F. & Lu, Y.-H. 2009. Customer churn prediction by hybrid neural networks. *Expert Systems with Applications,* 36, 12547-12553.

Van Exel, N. J. A. & Rietveld, P. 2009. Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travellers on the main travel corridors to the city of Amsterdam, The Netherlands. *Transportation Research Part A: Policy and Practice,* 43, 374-385.

van Lierop, D. & El-Geneidy, A. 2016. Enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit. *Research in Transportation Economics,* 59, 50-59.

Verplanken, B., Aarts, H. & Van Knippenberg, A. 1997. Habit, information acquisition, and the process of making travel mode choices. *European Journal of Social Psychology,* 27, 539-560.

Verplanken, B., Walker, I., Davis, A. & Jurasek, M. 2008. Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses. *Journal of Environmental Psychology,* 28, 121-127.

Vuchic, V. R. 2007. *Urban transit systems and technology,* Hoboken, N.J., Hoboken, N.J. : John Wiley & Sons.

Walker, I., Thomas, G. O. & Verplanken, B. 2015. Old Habits Die Hard. *Environment and Behavior,* 47**,** 1089-1106.

Walker, J. 2008. Purpose-driven public transport: creating a clear conversation about public transport goals. *Journal of Transport Geography,* 16**,** 436-442.

Webb, V. 2010. *Customer Loyalty in the Public Transportation Context.* Master of Science in Transportation, Massachusetts Institute of Technology.

Wieringa, J. E. & Verhoef, P. C. 2007. Understanding Customer Switching Behavior in a Liberalizing Service Market: An Exploratory Study. *Journal of Service Research : JSR,* 10**,** 174-186.

Yoh, A., Taylor, B. D. & Gahbauer, J. 2012. Does Transit Mean Business? Reconciling academic, organizational, and political perspectives on Reforming Transit Fare Policies. St. Louis: Federal Reserve Bank of St Louis.

Zhang, Y., Martens, K. & Long, Y. 2018. Revealing group travel behavior patterns with public transit smart card data. *Travel Behaviour and Society,* 10**,** 42-52.

Zhao, J., Qu, Q., Zhang, F., Xu, C. & Liu, S. 2017. Spatio-Temporal Analysis of Passenger Travel Patterns in Massive Smart Card Data. *IEEE Transactions on Intelligent Transportation Systems,* 18**,** 3135-3146.

Zhao, Y., Li, B., Liu, W. & Ren, S. 2005. Customer Churn Prediction Using Improved One-Class Support Vector Machine. *Advanced Data Mining and Applcations,* 3584**,** 300 - 306.

Zhao, Z., Koutsopoulos, H. N. & Zhao, J. 2018. Individual mobility prediction using transit smart card data. *Transportation Research Part C: Emerging Technologies,* 89**,** 19-34.

Zhong, C., Batty, M., Manley, E., Wang, J., Wang, Z., Chen, F. & Schmitt, G. 2016. Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data: e0149222. *PLoS ONE,* 11.

Part D: Appendices

Appendix A: Customer Fluctuation Survey Questionnaire

Note: This survey commences at Question 3 as the explanatory statement and participant consent were presented as Question 1 and Question 2.

Q3 Which gender do you identify as?

1. Male
2. Female
3. Other

Q4 Please select your age range?

1. under 18
2. 18 - 19
3. 20 - 29
4. 30 - 39
5. 40 - 49
6. 50 - 59
7. 60 – 69
8. 70+

Q5 What is your postcode? Open answer)

Q6 Can you please indicate which of the following best describes your current employment status?

1. Employed full-time
2. Employed part-time
3. Employed, Away from work
4. Unemployed looking for work
5. Student
6. Retired
7. Home duties
8. Other (provide answer)

Q7 Which of the following best describes your current household?

1. Family with children under 18
2. Family with adult children
3. Couple with no children
4. Single

5. Group
6. Other (provide answer)

Q8 Which bracket represents your personal weekly income? (Gross annual income in brackets)

1. Negative/nil income
2. $1 - $299 ($1 - $15,548 p.a)
3. $300-$499 ($15,600 - $25,948 p.a)
4. $500 - $799 ($26,000 - $41,548 p.a)
5. $800 - $1,249 ($41,600 - $64,948 p.a)
6. $1,250 - $1, 749 ($65,000 - $90,948 p.a)
7. $1,750 - $2,999 ($91,000 - $155,948 p.a)
8. $3000 or more ($156,000 or more p.a)
9. Prefer not to say

Q9 What is your highest educational qualification?

1. Post Graduate Degree
2. Bachelor Degree
3. Graduate diploma or certificate
4. Year 10 or above
5. No educational attainment

Q10 Did you use Public Transport within Melbourne at any time in the last year, November 2017 - October 2018?
Note: This includes one trip or more for the entire year

1. Yes
2. No

Following the completion of these questions, users were screened out or directed to the appropriate survey branch based on their response to Question 10. Users that answered no (2) to Question 10, were shown an additional question to identify if they were a lost user or non-user for screening out or completing the survey. This question was presented to participants as:

Q11 Did you use Public Transport within Melbourne in the previous year, November 2016 - October 2017?

1. Yes, I used the bus/train or tram regularly
2. Yes, I used the bus/train or tram sometimes
3. No, I never use public transport

Users that answered that they used public transport (bus/train or tram) regularly or sometimes were sent to a pre-identified “lost user” series of questions. This determined their primary mode of public transport and their reasons for stopping or temporarily ceasing their use of this mode. All users that answered “no, I never use public transport” were thanked for their time and screened out of the survey as ineligible.

For respondents that had used public transport in the previous year (regularly/sometimes), they were shown the following set of questions to identify the reasons behind why they had become a “lost” user.

Q12 When you use/d public transport in the previous year did you:

1. Mainly use the bus
2. Mainly use the tram
3. Mainly use the train

The next question was asked three times for each mode, but is provided once here to avoid undue repetition.

Q13/14/15 Please select the 3 most influential reasons you STOPPED or PAUSED catching the BUS/TRAIN/TRAM across all categories.

	Please select 3
Life Event	
I went on holiday	
I stopped studying/finished school	
My life circumstances changed (e.g. I had children, retired)	
Other	
Lifestyle	
I no longer needed to make trips with the bus/train/tram	
I was sick or injured and couldn't catch the bus/train/tram	
Catching the bus/train/tram was too difficult with children	
I bought a car	
I don't like the bus/train/tram	
Other	
The Bus/Train/Tram Service	
Catching the bus/train/tram takes too long	
Catching the bus/train/tram was too unreliable	
The bus/train/tram route I was using changed or stopped	
I felt uncomfortable on the bus/train/tram I was using (old, untidy, poor seats)	
The bus/train/tram wasn't available at the times I wanted to travel	
The bus/train/tram was too crowded	
Catching the bus/train/tram got too expensive	
Other	
The Trip Taken	
I travel at night and didn't feel safe	
The bus/train/tram wasn't needed as part of my journey anymore	
I found it easier to drive to my destination	
Other	

Q16 Do you have any other comments about how you make decisions to travel or not travel on public transport? (Open Answer)

The survey was then completed for this stream and participants were thanked for their time.

For participants that answered yes to Question 10, that they had travelled by public transport in the past year, they entered the public transport users stream.

Q17 When you use/d public transport in the last year did you?
Note: Please try to select the mode you use/d most to continue with this survey.

1. Mainly use the bus
2. Mainly use the tram
3. Mainly use the train
4. Mainly use more than one public transport mode (indistinguishable level of use)

Users that responded that they mainly used more than one public transport mode (4) were excluded from completing the remainder of the survey. Whereas, bus, train or tram users entered a mode specific set of questions about their trip taking behaviour. These streams will be shown once as the same questions as asked for all three modes. Where mode is displayed as Train/Tram/Bus, only the relevant mode was shown to participants. The questions for this stream were as follows:

Q18/26/34 Did you use the TRAIN/TRAM/BUS in the previous year, November 2016 - October 2017?

1. Yes, I used the train/tram/bus regularly
2. Probably, I use the train/tram/bus sometimes
3. No, I just started using the train/tram/bus this year

Q19/27/35 How often did you use the TRAIN/TRAM/BUS in the past month?

1. 5 - 7 days a week
2. 3 - 4 days a week
3. 1 - 2 days a week
4. At least once
5. I did not use the train/tram/bus in the past month

Q20/28/36 When you used the train/tram/bus:

1. Mainly use only the train/tram/bus
2. Very occasionally use tram/bus, train/tram, train/bus
3. Use tram/bus, train/tram, quite frequently but not as my main mode

Q21/29/37 For the last year, please select yes or no for whether you travelled by TRAIN/TRAM/BUS at least once for each month?

	Please select for each month	
	Yes (1)	No (2)

Nov '17		
Dec '17		
Jan '18		
Feb '18		
Mar '18		
Apr '18		
May '18		
Jun '18		
Jul '18		
Aug '18		
Sep '18		
Oct '18		

Q22/30/38 Using the statements below, how would you categorise your TRAIN/TRAM/BUS travel in the last year?

III. I started using the train/tram/bus a few months into the year and used it most months once I started

IV. I used the train /tram/bus for at least one month this year and then didn't use the train/tram/bus again

V. I used the train/tram/bus every month, or most months

VI. I used the train/tram/bus sometimes

These questions were used to direct users to the appropriate set of questions identifying their reasons for the applicable customer fluctuation behaviour. Those that answered that they started using the train/bus/tram a few months into the year, or those that used the train/bus/tram sometimes were directed to question 23/31/39.

Q23/31/39 Please select the top 3 most influential reasons you STARTED catching the TRAIN/TRAM/BUS this year or RETURNED to the TRAIN/TRAM/BUS after a break.

	Please Select 3
Life Event	
I changed home or work locations	
My life circumstances changed (I returned to work after having kids, I started a new job)	
I returned from holiday/ sabbatical	
Other	
Lifestyle	
I am trying to save money on my transport costs	
I was too sick or injured to drive or cycle	
I do not own a car or can't drive	
My car or bike was unavailable	
I believe it is important to take sustainable transport modes	
I have started a new hobby/ socialising more	
Other	
The Train/Bus/Tram Service	
I enjoy traveling by train/tram/bus	
I find catching the train/tram/bus reasonably priced	
I feel safest when catching the train/tram/bus	
I find the train/tram/bus to be less crowded than other modes	
I find the train/tram/bus to be reliable	
The train/tram/bus timetable changed to suit me better	
Other	
The Trip Taken	
I used the train/tram/bus when traveling to an event	
Parking is too difficult at my destination	
The train/tram/bus is the most convenient options for the main trips I was making	
I had an unpleasant experience with a different mode of transport	
I catch the train/tram/bus when the weather isn't suitable for other travel	
Other	

Those that answered that they used the train/bus/tram at least once and didn't use it again, or those that used the train/tram/bus sometimes were directed to question 24/32/40 around those that stopped or paused their train/tram/bus use.

Q24/32/40 Please select the 3 most influential reasons you STOPPED or PAUSED catching the TRAIN/BUS/TRAM across all categories.

	Please select 3
Life Event	
I went on holiday	
I stopped studying/finished school	
My life circumstances changed (e.g. I had children, retired)	
Other	
Lifestyle	
I no longer needed to make trips with the train/tram/bus	
I was sick or injured and couldn't catch the train/tram/bus	
Catching the train/tram/bus was too difficult with children	
I bought a car	
I don't like the train/tram/bus	
Other	
The Train/Tram/Bus Service	
Catching the train/tram/bus takes too long	
Catching the train/tram/bus was too unreliable	
The train/tram/bus route I was using changed or stopped	
I felt uncomfortable on the train/tram/bus I was using (old train, untidy, poor seats)	
The train/tram/bus wasn't available at the times I wanted to travel	
The train/tram/bus was too crowded	
Catching the train/tram/bus got too expensive	
Other	
The Trip Taken	
I travel at night and didn't feel safe	
The train/tram/bus wasn't needed as part of my journey anymore	
I found it easier to drive to my destination	
Other	

Finally, those that identified that they used the train/tram/bus regularly were directed to question 25/33/41 about why they regularly use their mainly used mode.

Q25/33/41 Please rank the 3 most influential reasons that you regularly catch the TRAIN/TRAM/BUS

	Please select 3
Life Event	
I changed home or work locations	
I started studying	
My circumstances haven't changed (e.g. I take my kids to school on the train/tram/bus)	
Other	
Lifestyle	
I believe it is important to take sustainable transport modes	
I don't have a car or can't drive	
I am trying to save on my transport expenses	
I was too sick or injured to drive	
I regularly use all modes of public transport	
Other	
The Train/Tram/Bus Service	
The train/tram/bus is the most convenient option for me	
I enjoy catching the train/tram/bus	
I feel comfortable catching the train/tram/bus	
I like that the train/tram/bus is not too crowded	
I find the train/tram/bus reliable	
I think the train/tram/bus provides a good service	
Other	
The Trip Taken	
I routinely make the same trip and like knowing what to do	
Parking is difficult at my regular destination	
I get the train/tram/bus as just one part of my regular journey	
Other	

Q26/34/42 Any other comments?