

This is unformatted version of paper: Zhang, C., Schmöcker, J.-D and Trépanier, M. (2022). Latent Stage Model for Carsharing Usage Frequency Estimation with Montreal Case Study. Transportation, 49(1), 185-211. <https://doi.org/10.1007/s11116-021-10173-9>

Latent Stage Model for Carsharing Usage Frequency Estimation with Montreal Case Study

Zhang Cen¹, Jan-Dirk Schmöcker¹ and Martin Trépanier²

¹Department of Urban Management, Kyoto University, Japan

²Department of Mathematical and Industrial Engineering and Interuniversity Research
Centre on Enterprise Networks, Logistics and Transport (CIRRELT), Polytechnique
Montréal, Canada

ABSTRACT

In order to predict the monthly usage frequency of members of a car-sharing scheme by analysing the gradual change of behaviour over time, a new model is proposed based on the Markov Chains model with latent stages. The model accounts for changing patterns of frequency from soon after signing up to later stages by including five latent user ‘life stages’. In applying the model to panel data from Montreal’s free-floating carsharing service the authors calculate each user’s ‘lifetime’ applied to ‘system operation time’, the time period since the start of the scheme. Three-fold validation reveals effective performance of the model for both lifetime and system operation time dimensions. The model is further applied to illustrate how previous carsharing experience and the extension of the scheme to a larger area can affect usage frequency changes. We conclude that this approach is effective for usage prediction for novel transport schemes.

Keywords: Carsharing, Usage frequency, User life cycle, Stochastic Process,
Hierarchical Hidden Markov Models

1. INTRODUCTION

Shared mobility, including carsharing, bicycle-sharing, micro transit, and on-demand ride services, has recently gained popularity as an alternative transport approach, especially in space-constrained urban areas where transportation costs are high. Carsharing can save costs, and space and reduce carbon emissions. According to Shaheen and Cohen (2016) nearly five million persons were registered for carsharing services worldwide as of 2014. In some cities and regions carsharing has been accepted fairly rapidly by local residents, whereas in others diffusion of the practice has lagged. Yoon et al. (2019), for example, noted that carsharing in China has been growing slower than in European or American cities, but that in Beijing it could potentially reach 9-32% of eligible users depending on a range of parameters. For usage prediction of new schemes we must consider their technological development as well as their popular acceptance including public promotion of the scheme (Stathopoulos et al., 2017).

Given these many contributing factors, analysis and prediction of travel demand for new mobility schemes remain challenging, but it is necessary in order for operators to create and revise service deployment and extension strategies. Operators may distinguish two major aspects affecting prediction of demand: Adoption, or how many new adopters will be attracted to the system; and behavioural dynamics, or how users in the system change their behaviour over time.

In this research, we focus on the second aspect, capturing the dynamics of usage frequency. Compared to traditional methods, such as time series analysis, stochastic process approaches that represent individual behaviour change may be more suitable to reflect these dynamics. Many researchers rely on Markov models to analyse panel data, but these models have a number of shortcomings (Zhang and Schmöcker, 2017). Zhang and Schmöcker have proposed an alternative approach that utilizes individual usage records over initial months in order to predict usage in subsequent months. They adopted the marketing concepts ‘customer life cycle’ and ‘ability and willingness to

purchase' to their framework in order to model user behaviour based on bicycle sharing data from Kyoto, Japan. This paper extends this approach by introducing a 5-stage life cycle model with latent stages that accounts for the existence of large differences in behavioural changes in the early and later time periods of the individual customer life cycle. We propose that this model achieves better forecasting performance and promotes a more accurate explanation of the different behaviour changes that occur over time, including the impacts of fluctuation in supply.

The paper applies the approach to data from a car-sharing scheme in Montreal. We demonstrate the impacts of a service area extension to the city's free-floating carsharing service (FFCs). We further estimate different stage and state transitions in usage frequencies for users of the FFCs scheme, depending on whether or not the person has prior experience with Montreal's station-based carsharing service (SBcs).

The paper is organized as follows: After this introduction, the authors will discuss adoption of carsharing services and will review problems associated with utilizing Markov models for user life cycle modelling, based on a literature review. In Section 3 we will describe the stochastic process of discrete behaviour observed at discrete time periods. Section 4 investigates how this process can be calibrated by observations using the maximum likelihood estimation. The authors propose use of an Expectation-Maximisation algorithm to accommodate the presence of latent variables in this model. Section 5 provides the Montreal case study to illustrate the feasibility of this approach and the above-mentioned applications. As part of the case study, we also demonstrate the validity of the model by applying three-fold validation and by using estimates from the first months to forecast usage in later time periods. We then summarize the major findings and discuss possible future research.

2. LITERATURE REVIEW

2.1 Adoption of carsharing services

There is an extensive and varied body of academic literature on carsharing usage and its impacts. Studies that analyse the impacts of carsharing have identified multiple benefits of these schemes, such as reduced car ownership rates in urban areas, lower demand for parking space, reduced vehicle mileage leading to lower greenhouse gas emission levels, heightened physical activity, individual financial savings and improved multimodal transport connections (Cervero and Tsai, 2004; Martin and Shaheen, 2011). Such studies have been conducted based on observed trip data as well as agent-based simulation (Balać et al., 2015; Martinez et al., 2017). Other studies have analysed member characteristics based on usage data, resulting in a range of findings. There is broad consensus that carsharing users tend to be younger and more educated than the average population (Burkhardt and Millard-Ball, 2006; Efthymiou et al., 2013). Shaheen and Schwartz (2004) found that carsharing users are often students and belong to low-income households. Coll et al. (2014) concluded that socio-economic profiles, education, income, family structure, and nonmotorized mode usage are strong predictors of carsharing membership. Some studies further suggested that carsharing members are more environmentally conscious (Costain et al., 2012; Clewlow, 2016).

Researchers on carsharing have applied a range of methodologies in order to estimate usage uptake or frequency estimation. These include many examples of stated preference surveys evaluating the importance of attributes such as cost and socio-demographic attributes but also variable factors such as weather (Catalano et al., 2008; Ciari and Axhausen, 2013; Yoon et al., 2017). Qualitative research, such as that by Burkhardt and Millard-Ball (2006), has been conducted implementing focus group interviews. These have also tended to reveal attitudinal differences among carsharing users.

We note that the majority of the above-mentioned studies were based on data or surveys regarding station-based one-way carsharing systems rather than free-floating schemes, despite the increasing popularity of the latter. The omnipresence of smartphones and other portable internet-connected devices has made free-floating systems possible, enabling users to easily locate available cars in their vicinity. This can mean an increase in user convenience and usage uptake if the vehicle density is large enough, as has been shown with agent-based models (Ciari et al., 2015; Heilig et al., 2015). Our research thus focuses on free-floating carsharing.

Several recent studies have examined usage evolution by introducing Rogers' technology diffusion theory. Shaheen et al. (2009) applied the theory using a system perspective to distinguish three phases in the deployment of carsharing schemes: initial market entry and experimentation, growth and market diversification, and finally commercial mainstreaming.

Several key factors influencing early adoption have been analysed in the literature. Shaheen et al. (1999) conducted a longitudinal survey of individuals interested in joining a carsharing program and found that sociodemographic (age, gender, income, auto ownership) and psychographic characteristics (attitudes toward current transport modes, vehicles, congestion, environment, and experimentation) impact an individual's decision to participate. Meijkamp (1998) categorized the possible determinants of carsharing adoption as personal attributes (e.g., car ownership, auto use frequency), service-oriented determinants (e.g., whether carsharing is available near one's home) and context-oriented determinants (e.g., rising vehicle costs, fuel prices). Namazu et al. (2018) wrote that the population of carsharing users is also related to the operational 'stage' of the carsharing system itself, as they found significant differences in the sociodemographic characteristics of those adopting carsharing when a scheme is initiated and those who adopt it at later stages.

These research papers share a common focus on understanding the role, use, impacts and adoption of carsharing service. However, we suggest there is still limited knowledge about the temporal evolution of individual usage frequency within the initial years, a period when carsharing systems often also undergo rapid changes in supply, as in the establishment of new stations or, in the case of FFcs, additions to the existing fleet. Recent studies are mostly based on surveys of carsharing users due to a lack of accurate trip datasets over long periods and suitable analysis methods. This study seeks to minimize this gap by analysing longitudinal data from Montreal, Canada. Wielinski et al. (2019) also analysed Montreal data, finding an increase in higher usage intensity classes and a change in the demographic composition of new members adopting the service over time. The current paper extends Wielinski's by developing a descriptive model for 'frequency transitions' of individuals in the carsharing population that we then apply to estimate and forecast system demand.

2.2 Markov Models in the Transport Literature

Markov models are used to describe these frequency transitions, assuming, in their basic form, that future states depend only on the current state and that transition probabilities stay constant over time (Kalbfleisch and Lawless, 1985). Markov chains have found wide application in the social sciences, especially in the study of data that record life history events for individuals (Bartholomew, 1982; Singer, 1981). In these studies, methods were developed to infer the parameters of a Markov process based on a matrix of transition probabilities between states, estimated on data from discrete time points.

A significant body of literature related to transportation demand forecasting has addressed different applications of Markov models and general stochastic process modelling, particularly focusing on day-to-day dynamics (Watling, 1996; Zhu et al., 2016). Zhang and Schmöcker (2017) discussed the potential issues arising from these

approaches in more detail, noting that most of this literature examines the circumstances under which (multiple) equilibria occur, as well as supply and demand interaction dynamics. In contrast, the research presented here focusses on the estimation of the transition matrix parameters of such a stochastic process.

In many cases, the events of interest here may not be directly observable. Hidden Markov models allow us to define both observed events and hidden events that may be causal factors in a probabilistic model. Such models are often applied in reinforcement learning and temporal pattern recognition, such as with speech, handwriting and gesture recognition. They are also used for a range of applications in the transport literature, such as describing and evaluating network traffic flows (Liu et al., 2015), characterizing activity sequences and imputing trip purposes (Han and Sohn, 2016), detecting driving fatigue (Rashwan et al., 2013) and modelling driving behaviour (Tang et al., 2015). In a study with similar objectives to the research contained here, Xiong et al. (2015) used a Hidden Markov approach to estimate changes in mode choice between different waves of a panel survey. The hidden stages in their model reflected long-term user characteristics such as family status. Their model was derived from Bayesian estimation and Markov Chain Monte Carlo simulation. Yang et al. (2017) and Xiong et al. (2018) applied and extended the model to describe the dynamics associated with car ownership decisions. In contrast to this approach, however, in the current research membership in a ‘life stage’ is not inferred from sociodemographic data. Instead, the dataset is based on a series of usage records with limited additional information. As will be explained in the following section, the hidden stages in our approach simply define a progression in the usage history similar to life cycles used in marketing literature.

2.3 Previous Research on Life Cycle Models

According to customer relationship marketing literature (e.g. Kumar and Reinartz, 2018), the customer life cycle comprises key milestones that every customer achieves over the course of a relationship with a brand. For membership schemes customers sign

up, make their first purchase and then some turn into repeat buyers. At some point, however, they will reduce and then stop purchasing the product due to changes in personal circumstances or product changes. We suggest application of this concept, which typically comprises five stages according to Cox (1967), in order to understand the behaviour of a population exploring new transport options, as shown below.

1. Birth stage (B): An individual becomes aware of the new transport service and tries to use it. “Birth” denotes use of this system for the first time. All individuals start at this stage.
2. Growing stage (G): The traveller becomes increasingly familiar with the transport system and is likely to increase travel frequency.
3. Maturity stage (M): The user has finished the adoption process and hence the travel frequency remains fairly constant.
4. Recession stage (R): Travel frequency reduces over time as, for example, another transport service starts to become a better option.
5. Death stage (D): The traveller stops using the service and is unlikely to return. This is defined as an absorbing stage.

This five-stage model is an extension of Zhang and Schmöcker’s (2017) model, which only considered three stages: ‘Birth’, ‘Maturity’ and ‘Death’. In Zhang and Schmöcker’s initial model Birth was defined as the first month of usage and Death as the last, denoting stage membership yet not requiring definition of probabilistic latent stages. However, this simplification may have reduced the accuracy of the model. We found that significant differences in transition matrices between initial and final time periods existed when applying the model to real data. In this paper we suggest that latent stages of varying duration for users can provide better model fits. In the following section we describe this approach of simultaneous stage and state estimation.

3. MODEL CONCEPTUALIZATION

3.1 Notation

The following notation will be used throughout the paper. We group the variables according to whether they are time-dependent or not. Further, variables marked with a tilde accent refer to estimated variables, applying a maximum likelihood approach as explained in Section 4 to best reflect (and forecast) available data.

Table 1 Notation

Type	Variable	Description
Time-Independent Variables	T	Discrete system operation time periods since the operation of the system with $t=0, 1, 2, \dots$
	τ	Discrete individual lifetime periods since initial usage with $\tau=0, 1, 2, \dots$
	J	Set of users recorded in the test data set with j denoting a single user
	L	Set of life cycle stages $L=(B,G,M,R,D)$ with indices l and s used to denote the possible life stage of a user at a given time
	$\pi [\pi_{ls}]$	Stage transition matrix with elements π_{ls} denoting the transition probability from stage l to stage s
	S	Set of usage states
	m	The number of usage states in set S . In our case study, $m=6$
	Φ, Φ_l	Set of stage-specific state transition matrices. $\Phi=(\Phi_B, \Phi_G, \Phi_M, \Phi_R, \Phi_D)^T$
Time-Dependent Variables	$\Phi_l[\phi_{ik}^l]$	State transition matrix for users in stage l with element ϕ_{ki}^l denoting the state probability for users in stage l to change from state i to state k
	$b(t)$	Number of persons ‘being born’, that is entering the system at time t (Input variable)
	$X(t)$	Observed system state vector at time t as aggregation of vector $x_j(t)$
	$\Delta X(t)$	Observed state aggregation vector for new adopters $b(t)$ at time t
Time-Dependent Variables	$x_j(t)[x_j^i(t)]$	Observed state vector of person j at time t with binary elements $x_j^i(t)$ denoting whether or not person j is in state i at time t (binary)

$\tilde{\mathbf{X}}(t)$	Estimated system state at time t as aggregation of vector $\tilde{\mathbf{x}}_j(t)$
$\tilde{\mathbf{x}}_j(t)[\tilde{x}_j^i(t)]$	Estimated state probability distribution of user j at time t with elements $\tilde{x}_j^i(t)$ denoting the estimated probability of person j being in state i at time t
$\tilde{\mathbf{Q}}(t)$	Estimated system stage vector at time t as the aggregation of $\tilde{\mathbf{q}}_j(t)$
$\tilde{\mathbf{q}}_j(t)[\tilde{q}_j^l(t)]$	Estimated stage probability distribution vector of person j at time t with elements $\tilde{q}_j^l(t)$ denoting the probability of user j being in a stage l at time t
$n_i(\tau), \tilde{n}_i(\tau)$	Observed/Estimated number of persons in state i at life time τ
$n_{ik}(\tau - 1, \tau), \tilde{n}_{ik}(\tau - 1, \tau)$	Observed/Estimated number of persons transferring from state i at time $\tau - 1$ to state k at life time τ

3.2 Description of Lifecycle Model

We denote the discrete-time periods since the initiation of the system with $t = 0, 1, 2, \dots$. Suppose the process of individual demand change over time t can be represented by a set of discrete states S with m states as $\{1, \dots, m\} \in S$. These discrete states can represent behaviour such as usage frequency, spatial distribution, or other measurements, according to research requirements. In this case study, we will use $m=6$ with usage states 1 to 6 denoting usage frequency from low to high (see Section 5.1). For user j a transition from one state to another can take place between adjacent time periods t and $t+1$. Since it is a Markov process, these transitions are probabilistic. The Markov property implies that the probability of moving to the next state depends only on the present state. The process is further ‘time-homogeneous’, meaning that the state transitions are independent of any ‘absolute time’ measure such as the season and the length of time a user has been registered.

Our motivation for the introduction of a life-cycle model stems from the fact that the time-homogeneity assumption, in particular independence with respect to usage period, does not always hold. To capture this, we introduce different transition matrices to fit

the usage dynamics. We denote discrete individual lifetime periods since initial usage with $\tau = 0, 1, 2 \dots$ and a set of life stages Birth (B), Growing (G), Maturity (M), Recession (R) and Death (D) with $\{B, G, M, R, D\} \in L$. In the following, for brevity, we refer to ‘life stage’ often as ‘stage’ and ‘usage state’ as ‘state’. We further define the matrix π for life stage transitions and the matrices $\phi_B, \phi_G, \phi_M, \phi_R, \phi_D$ for usage state transitions of users in the respective life stage. Once these six matrices are Established we can obtain a time series of estimated usage frequencies. The relationship between ‘state’ and ‘stage’ can be regarded as analogous to the relationship between ‘speed’ and ‘acceleration’ for a driver who only observes the car dashboard. ‘State’ is then like ‘speed’ in that it represents the observable, current situation and information possessed by the driver, whereas ‘stage’ is like ‘acceleration’, in that the driver cannot directly measure it but it determines future speed changes. Furthermore the same speed is obtainable for different degrees of acceleration.

A schematic diagram of the 5-stage life cycle model is shown in Figure 1, using a directed graph to depict the stage and state transitions. In accordance with the life stage concept, we presuppose that a user can stay or progress to a later stage but cannot move backward.

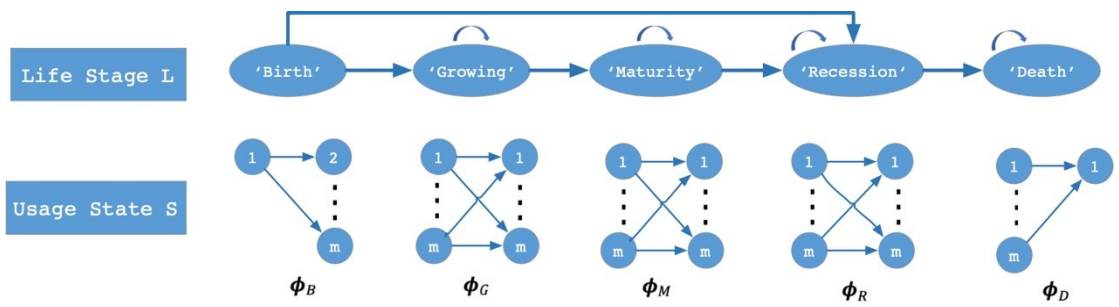


Figure 1 The 5-Stage Life Cycle Model

Further, this model does not permit transitions $\pi_{BB}, \pi_{BM}, \pi_{BD}, \pi_{GR}, \pi_{GD}, \pi_{MD}$. The resulting transition matrix π is shown in Table 2. Our rationale for the additional restrictions are as follows: First of all, we want stage B to represent the initial trial of

the new service, so we remove π_{BB} to ensure that users cannot stay at this stage after entering. Further, stage R should reflect the pattern of declining usage before departing the scheme. To ensure that all users go through R before D, we remove $\pi_{BD}, \pi_{GD}, \pi_{MD}$. Furthermore, stage G should reflect the ‘adaptation’ process after entering, whereas stage M should reflect the longer term usage pattern. ‘Adaptation’ in this study refers to the gradual formation of regular usage patterns, whether high or low utilization. We also note that we cannot guarantee the uniqueness of G and M if we include π_{BM} and π_{GR} . For example, for stage G the forward stages are G, M, R and the backward stages are B and G. We denote this as G: (B,G)-(G,M,R). Similarly, we can also denote M as M: (B,G,M)-(M,R). Assuming that G, M are an identical stage X, the model can be simplified to G: (B,X) - (X,R) and M: (B,X) - (X,R), so that we cannot reject the hypothesis that these stages are the same.

These simplifications reduce the number of parameters to be estimated. The life stage transition matrix now contains only eight stage transitions, as shown in Table 2. Since a user must be placed in one stage we have added a requirement that the probabilities in each row must add up to one ($\pi_{BG} + \pi_{BR} = 1, \pi_{GG} + \pi_{GM} = 1, \pi_{MM} + \pi_{MR} = 1, \pi_{RR} + \pi_{RD} = 1$) so that only four parameters remain. The first parameter in each row is selected for estimation, i.e. $\pi_{BG}, \pi_{GG}, \pi_{MM}, \pi_{RR}$. Our interpretation of these parameters is as follows:

π_{BG} represents the percentage of potential long-term users, i.e. those who have a positive attitude to the scheme and hence will go through the whole life cycle process. In contrast, $1 - \pi_{BG}$ is the probability of a user finding the scheme unsatisfying and after trial proceeding directly to the recession stage R. $\pi_{GG}, \pi_{MM}, \pi_{RR}$ can reflect the length of the adaptation process. π_{GG} indicates the period over which users keep increasing their usage frequency, π_{MM} the time over which users remain system members without a tendency to drop out and π_{RR} the period over which we generally accept declining usage until users eventually leave the scheme.

Table 2 Transition Matrix π for Life Cycle

STAGE	B	G	M	R	D
B	-	π_{BG}	-	$\pi_{BR}=1-\pi_{BG}$	-
G	-	π_{GG}	$\pi_{GM} = 1 - \pi_{GG}$	-	-
M	-	-	π_{MM}	$\pi_{MR} = 1 - \pi_{MM}$	-
R	-	-	-	π_{RR}	$\pi_{RD}=1-\pi_{RR}$
D	-	-	-	-	1

There are also some restrictions in usage state transitions for those in life stages B and D. For Stage B, before using the service, there is clearly no usage so that the matrix reduces to a vector with transitions from a no-usage state (State 1 in the case study) to any other state. Similarly, in life stage D users will not use the service anymore so that there is only a vector of transitions from any state to the no-usage state. There are no restrictions in Stages G, M and R.

3.3 Formulation of Life Cycle Model

Let $\mathbf{X}(t)$ be the observed system state and $\mathbf{x}_j(t) = (x_j^1(t), \dots, x_j^m(t))$ be the person or person group j specific usage frequency vector state at time t with elements $x_j^i(t) \subseteq (0,1)$ denoting whether a person j is in state i or not. $\mathbf{X}(t)$ can be represented by aggregation of $\mathbf{x}_j(t)$, i.e. $\mathbf{X}(t) = \sum_{j \in J} \mathbf{x}_j(t)$. Each person must be in exactly one state at each time t so that $x_j^i(t)$ takes binary values with constraint $\sum_{i=1 \dots m} x_j^i(t) = 1$.

Unlike with observable usage states, life stages are latent variables. Let $\tilde{\mathbf{Q}}(t)$ be stages of the system and $\tilde{\mathbf{q}}_j(t)$ and vector $\tilde{q}_j^l(t)$ be the estimated probability mass function for user j which we refer to as the ‘stage probability distribution’. $\tilde{q}_j^l(t)$ is the estimated probability of user j being in stage l at time t .

We are interested in estimating the transition probabilities between subsequent time

periods. For this, we define $\tilde{X}(t)$, $\tilde{x}_j(t)$ and $\tilde{x}_j^i(t)$ as estimated state variables for $X(t)$, $x_j(t)$ and $x_j^i(t)$. It is clear that for our estimated parameters constraints $\sum_{i=1\dots m} \tilde{x}_j^i(t) = 1$ and $\sum_{l=B,G,M,R,D} \tilde{q}_j^l(t) = 1$ must hold and $\tilde{x}_j^i(t)$ and $\tilde{q}_j^l(t)$ are bounded by 0 and 1. Due to potential sampling limitations, including potential errors and discretization of the states, we do not include the boundaries.

We further define the Markovian transition function $\boldsymbol{\pi}$ in stages and $\boldsymbol{\phi}_l$ in states for different stages l , which represent the transition probabilities of the processes to update the estimated stage and state probabilities. We define $\boldsymbol{\pi}$ as a 5×5 matrix denoting the probability of stage transitions between stages B to D. For different stages, the forms of matrices $\boldsymbol{\phi}_l$ are different. $\boldsymbol{\phi}_B$ is a $1 \times m$ vector for new ‘entering’ users since before initial usage, all people are in State 1 (no usage). $\boldsymbol{\phi}_G, \boldsymbol{\phi}_M, \boldsymbol{\phi}_R$ are $m \times m$ matrices denoting the probability of state transitions between states 1 to m . Finally, $\boldsymbol{\phi}_D$ is a $m \times 1$ vector for departed users since the usage state of all users in this state must be ‘inactive’. We denote all elements of the transition matrices with ϕ_{ik}^l where i and k range from states 1 to m . Similarly, we denote $\boldsymbol{\phi} = (\boldsymbol{\phi}_B, \boldsymbol{\phi}_G, \boldsymbol{\phi}_M, \boldsymbol{\phi}_R, \boldsymbol{\phi}_D)^T$ as the stage transition matrices (we extend $\boldsymbol{\phi}_B, \boldsymbol{\phi}_D$ into $m \times m$ matrices with adding 0).

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_{BB} & \cdots & \pi_{BD} \\ \vdots & \ddots & \vdots \\ \pi_{DB} & \cdots & \pi_{DD} \end{bmatrix} \quad (1)$$

$$\boldsymbol{\phi}_l = \begin{bmatrix} \phi_{11}^l & \cdots & \phi_{1m}^l \\ \vdots & \ddots & \vdots \\ \phi_{m1}^l & \cdots & \phi_{mm}^l \end{bmatrix} \quad (l = B, G, M, R, D) \quad (2)$$

3.4 Properties of the Life Cycle Model

We assume that $Q(t)$, $t > 0$ (the stage transition) is a stochastic process, the transition is time-homogeneous and the Markov property holds. The Markov assumption also applies for the state transitions $X(t)$, $t > 0$, except that the state transition is time-

homogeneous only when in the same stage, and the probability of moving to the next state depends on the present stage and state. Under this assumption, for user j in stage l we can update the state and stages as in (3) and (4).

$$\tilde{\mathbf{q}}_j(t) = \tilde{\mathbf{q}}_j(t - r) \cdot \boldsymbol{\pi}^r \quad (3)$$

$$\tilde{\mathbf{x}}_j(t) = \mathbf{x}_j(t - 1) \cdot \sum_l (\tilde{\mathbf{q}}_j^l(t - 1) \cdot \boldsymbol{\phi}_l) \quad (4)$$

In (3) we obtain the stage distribution t by applying the stage transition matrix $\boldsymbol{\pi}$ to the stage distribution in time period r . Note that we can obtain the expected stage distribution from an initial starting point at time 0. In (4) the summation term weighs the state transition matrices by the likelihood of being in a state. For this summation of the entry-wise product of $\tilde{\mathbf{q}}_j$ with $\boldsymbol{\phi}_l$ we introduce the shorthand $\sum_l (\tilde{\mathbf{q}}_j^l(t) \cdot \boldsymbol{\phi}_l) = \tilde{\mathbf{q}}_j(t) \star \boldsymbol{\Phi}$. Following our previous discussion, we further distinguish operation-time periods and user lifetime periods. By denoting b_j as the time period in which a user is ‘born’, i.e. first using the scheme, and τ as individual lifetime period since ‘birth’, we further construct $\tilde{\mathbf{q}}_j(\tau) = \tilde{\mathbf{q}}_j(t - b_j)$, $\mathbf{x}_j(\tau) = \mathbf{x}_j(t - b_j)$, $\tilde{\mathbf{x}}_j(\tau) = \tilde{\mathbf{x}}_j(t - b_j)$ to denote the estimated life stage and observed/estimated usage state in the user specific lifetime. Combining (3) and (4) using this shorthand, the person-specific state probability distribution $\tilde{\mathbf{x}}_j(\tau)$ can now be estimated by $\mathbf{x}_j(t)$ and transition matrices $\boldsymbol{\pi}$ and $\boldsymbol{\Phi}$:

$$\tilde{\mathbf{x}}_j(\tau) = \mathbf{x}_j(\tau - 1) \cdot \left((\tilde{\mathbf{q}}_j(\tau = 0) \cdot \boldsymbol{\pi}^{\tau-1}) \star \boldsymbol{\Phi} \right) \quad (5)$$

In the ‘individual lifetime’ domain, all users start at the same time period in stage B at $\tau = 0$. Therefore for all users, $\tilde{\mathbf{q}}_j(\tau = 0) = (1, 0, 0, 0, 0)$ holds, and we denote this as $\mathbf{q}_B$. The estimated state distribution of the system for users in lifetime τ , denoted as $\tilde{\mathbf{X}}(\tau)$ can now be updated as in (6):

$$\tilde{\mathbf{X}}(\tau) = \sum_j \tilde{\mathbf{x}}_j(\tau) = \mathbf{X}(\tau - 1) \cdot \left((\mathbf{q}_B \cdot \boldsymbol{\pi}^{\tau-1}) \star \boldsymbol{\Phi} \right) \quad (6)$$

In the ‘system operation time’ domain, since there are new adopters, the population is dynamic. To account for such population dynamics we introduce a vector $\Delta\mathbf{X}(t-1)$ denoting state aggregation of $\mathbf{x}_j(t)$ for new adopters $b(t)$ before adoption as in (7).

$$\Delta\mathbf{X}^T(t-1) = (b(t), 0, \dots, 0) \quad (7)$$

With this, finally, the estimated system state $\tilde{X}(t)$ can be calculated as in (8) where $j(t)$ denotes the users recorded at time period t and $\tilde{\mathbf{Q}}(t-1)$ the estimated stage distribution of all users in the previous time period.

$$\tilde{X}(t) = \sum_{j(t)} \tilde{\mathbf{x}}_j(t) = \mathbf{X}(t-1) \cdot (\tilde{\mathbf{Q}}(t-1) \cdot \boldsymbol{\pi} \star \boldsymbol{\Phi}) + \Delta\mathbf{X}(t-1) \cdot (\mathbf{q}_B \star \boldsymbol{\Phi}) \quad (8)$$

4. MODEL ESTIMATION

The observed states $\mathbf{x}_j(t)$ for users at different time periods can be identified from usage data. From the first usage, we can further identify the ‘birth’ stage $\tilde{q}_j^B(t)$, but identifying all other stages including ‘death’ is not as straightforward. Last usage does not necessarily mean ‘death’, as the user might be only temporarily inactive and return to the scheme at a future time after the end of our observation period. In Zhang and Schmöcker (2017) we derived a probabilistic approach to classify inactive users as departed users. However, in the present approach we also treat stage D as latent.

We estimate the Markovian transition functions $\boldsymbol{\Phi}$ for states and $\boldsymbol{\pi}$ for stages. The objective is to maximize the likelihood L^* of correctly predicting the number of state transitions in each time period by using one step transition probabilities. We can formulate this as in (10) where $\tilde{n}_{ik}(\tau)$ and $n_{ik}(\tau)$ denote the estimated and observed number of persons transferring from state i in time period $\tau-1$ to state k in time period τ . $n_{ik}(\tau)$ is obtained as in (9) and $\tilde{n}_{ik}(\tau)$ is obtained accordingly with $\tilde{x}_j^i(\cdot)$ instead.

$$n_{ik}(\tau) = \sum_j x_j^i(\tau-1) x_j^k(\tau) \quad (9)$$

$$L^*(\boldsymbol{\Phi}, \boldsymbol{\pi}) = \prod_{\tau} \prod_{i \in M} \prod_{k \in M} (\tilde{n}_{ik}(\tau))^{n_{ik}(\tau)} \quad (10)$$

As we want to maximize the likelihood, we consider the log-likelihood function L:

$$\arg \max_{\boldsymbol{\phi}, \boldsymbol{\pi}} L = \arg \max_{\boldsymbol{\phi}, \boldsymbol{\pi}} \sum_{j \in J} \sum_{\tau} \sum_{i \in M} n_{ik}(\tau - 1, \tau) \ln \tilde{n}_{ik}(\tau) \quad (11)$$

Expanding (5) we can obtain the estimated term in (10) with (11) where we remind the reader that following (3) we can obtain $\tilde{\mathbf{q}}_j(t)$ recursively.

$$\tilde{n}_{ik}(\tau) = \sum_l \sum_s n_i(\tau - 1) \cdot (\tilde{q}_j^l(\tau - 1) \cdot \pi_{ls}) \cdot \phi_{ik}^l \quad (12)$$

The resulting function is shown in (13) with constraints (14) and (15).

$$\begin{aligned} & \arg \max_{\boldsymbol{\phi}, \boldsymbol{\pi}} L(\boldsymbol{\phi}, \boldsymbol{\pi}) \\ &= \arg \max_{\boldsymbol{\phi}, \boldsymbol{\pi}} \sum_{\tau} \sum_i \sum_k n_{ik}(\tau - 1, \tau) \ln(\sum_l \sum_s n_i(\tau - 1) \cdot (\tilde{q}_j^l(\tau - 1) \cdot \pi_{ls}) \cdot \phi_{ik}^l) \end{aligned} \quad (13)$$

$$\sum_l \pi_{ls} = 1, \sum_{i=1 \dots m} \phi_{ik}^l = 1 \quad \forall k, s \quad (14)$$

$$0 < \phi_{ik}^l < 1, 0 < \pi_{ls} < 1 \quad (15)$$

We can identify the observed states $\mathbf{x}_j(\tau)$, and, then directly estimate the ‘birth’ stage $\tilde{q}_j^B(\tau)$. $\tilde{q}_j^G(\tau)$, $\tilde{q}_j^M(\tau)$, $\tilde{q}_j^R(\tau)$, and $\tilde{q}_j^D(\tau)$ are latent variables and need to be estimated together with the transition probabilities. Therefore the equations cannot be solved by known data observations directly. Here an expectation-maximization (EM) algorithm as discussed in Dellaert (2002) and Bishop (2006) is used to estimate the parameters and latent variables in our models. The EM algorithm seeks to find the Maximum Likelihood Estimation (MLE) of the marginal likelihood by iteratively applying the Expectation step (E) and the Maximization step (M): In Step E the expected value of the log-likelihood function is calculated with respect to the conditional distribution of latent variables under the current estimate of the parameters. In Step M one finds the parameters that maximize this quantity. This leads to an iterative algorithm to determine $\boldsymbol{\pi}$ and $\boldsymbol{\phi}$ as explained below.

1. Initialize the parameters in $\boldsymbol{\pi}$
2. Compute the probability of each possible value of $\boldsymbol{\phi}$ given $\boldsymbol{\pi}$

3. Use the computed ϕ values to obtain a better estimate for the parameters in π
4. Iterate Steps 2 and 3 until convergence

We note that convergence to a global optimum cannot be guaranteed. In the following case study we therefore conducted tests with different initial parameter settings but did not find convergence to different local optima to be an important issue for reasonable initial estimates.

5. CASE STUDY

5.1 The Communauto Carsharing System

Communauto, the oldest carsharing operator in North America, has provided a station-based carsharing service (SBcs) in Montreal since 1995. In 2013 the company introduced a new free-floating carsharing service (FFcs) called ‘Auto-mobile’. Table 3 and Figure 2 display the five extensions of the scheme that took place between July 2013 and April 2017. As the service area expanded, the car fleet was increased. As of April 2017, about 543 vehicles were part of the FFcs car fleet while the SBcs service comprised over 1030 vehicles at a total of 413 stations.

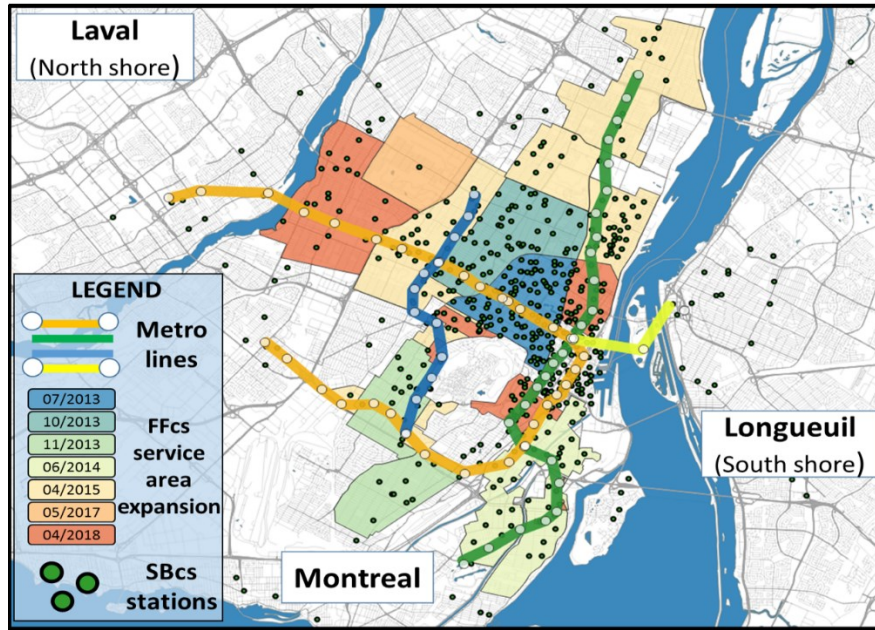


Figure 2 The “Auto-Mobile” FFcs service in Montreal (Wielinski, 2019)

Table 3 FFcs service area extensions in Montreal considered in this study

Service Areas	Start of operation (months since start of operation)		Area (km ²)	Active Vehicles
1 PLT	Jul-13	(1)	8.0	24
2 RSMT	Oct-13	(4)	18.1	24
3 CDN	Nov-13	(5)	30.2	64
4 SO	Jun-14	(12)	43.1	102-182
5 EST	April-15	(22)	78.1	215-543

Figure 3 shows the total usage and new adopters of the system over time. One can observe that the total usage growth has been fairly steady but that the number of people beginning to use the service recorded a distinctive peak in month 25 which follows the fifth extension of the service area (see Figure 2). There are more than 1.2 million individual FFcs usage records with over 15,000 users for the first 47 months of operation since July 2013. The data we obtained also contains registration date for users of both the FFcs and SBcs service in Montreal.

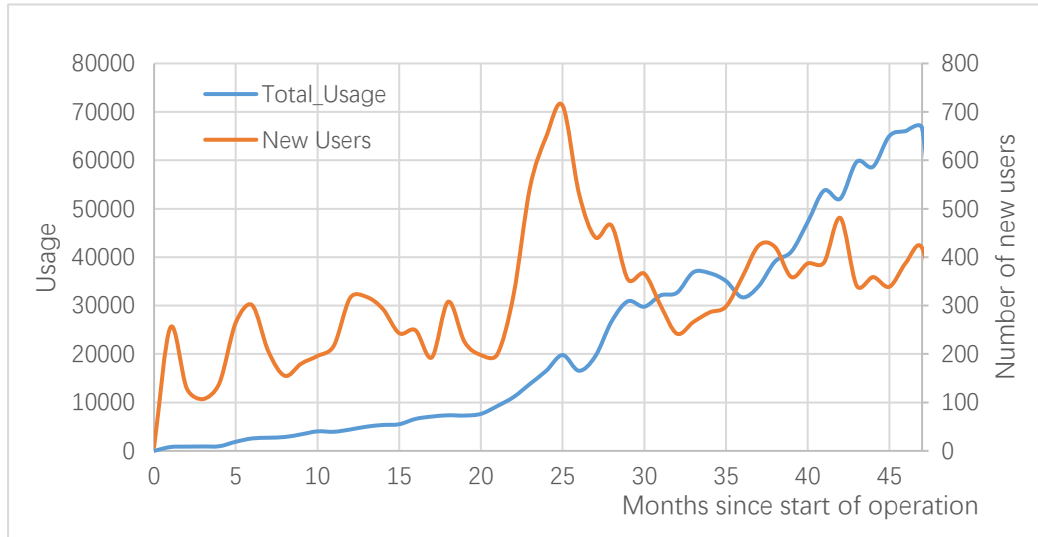


Figure 3 New users (initial usage month) and usage of the system over time

We distinguish six different states for each user according to use frequency and, as a reference point, show the number of users in all states in May 2017 (see Table 4). The rationale for cut-off points is that we distinguish those being inactive, those making at most one round trip per month (low), those making occasional trips (low-medium), those using the scheme fairly regularly (medium), those using the scheme, on average, at least every other day (medium-high) and those using the scheme, on average, every day (high). We note that, even though the exact boundaries for states are disputable, we tested sensitivities and found that small shifts would not significantly alter our data analysis and conclusions due to a fairly continuous distribution of usage frequencies. For example, if we change the boundaries for the medium states by ± 1 the estimation values of all stage transitions and most state transition parameters reported in section 5.3 change by less than 0.05 and we do not find differences in the overall conclusions. The time intervals between subsequent time points are set to one month.

Table 4 Usage State Divisions and Number of Users

	States	Usage Frequency (times in one month)	Users in May 2017
1	Inactive	0	9174
2	Low	1-2	2206
3	Low-medium	3-6	1636
4	Medium	7-15	1060

5	Medium-high	16-30	685
6	High	>30	662

5.2 Validation in Lifetime

Here we estimate the parameters of this 5-stage life cycle model with data from the first 40 months of operation. The estimations reported below were carried out using Python programming on a standard PC and were completed within 20-40 minutes depending on data sets and initial parameter settings.

To validate the reliability of the estimation results in user lifetime, the data for the first 40 months was evenly split into 4 data sets. Data sets 1- 3, as training sets, were used for three-fold cross-validation to avoid overfitting. We carried out three estimations. In each estimation two data sets were used to train the model and the remaining one was used to test the estimation performance. The average values of parameters in the three estimations were taken as the model parameters. Data set 4 was used as an out-sample test to check the performance of the model. In sum, we had four tests; three in-sample tests and one out-sample test.

To evaluate the model goodness of fit let $E_{i,k}(t)$ and $A_{i,k}(t)$ represent the estimated and actual proportion of users at lifetime t respectively in test k . The RMSPE (Root Mean Square Percentage Error) can then be calculated by (16).

$$R(k) = \frac{\sqrt{\sum_{t=1}^T \sum_{i=1}^m (E_{i,k}(t) - A_{i,k}(t))^2 / (m \cdot T)}}{\sum_{t=1}^T \sum_{i=1}^m A_{i,k}(t) / (m \cdot T)} \times 100\% \quad (16)$$

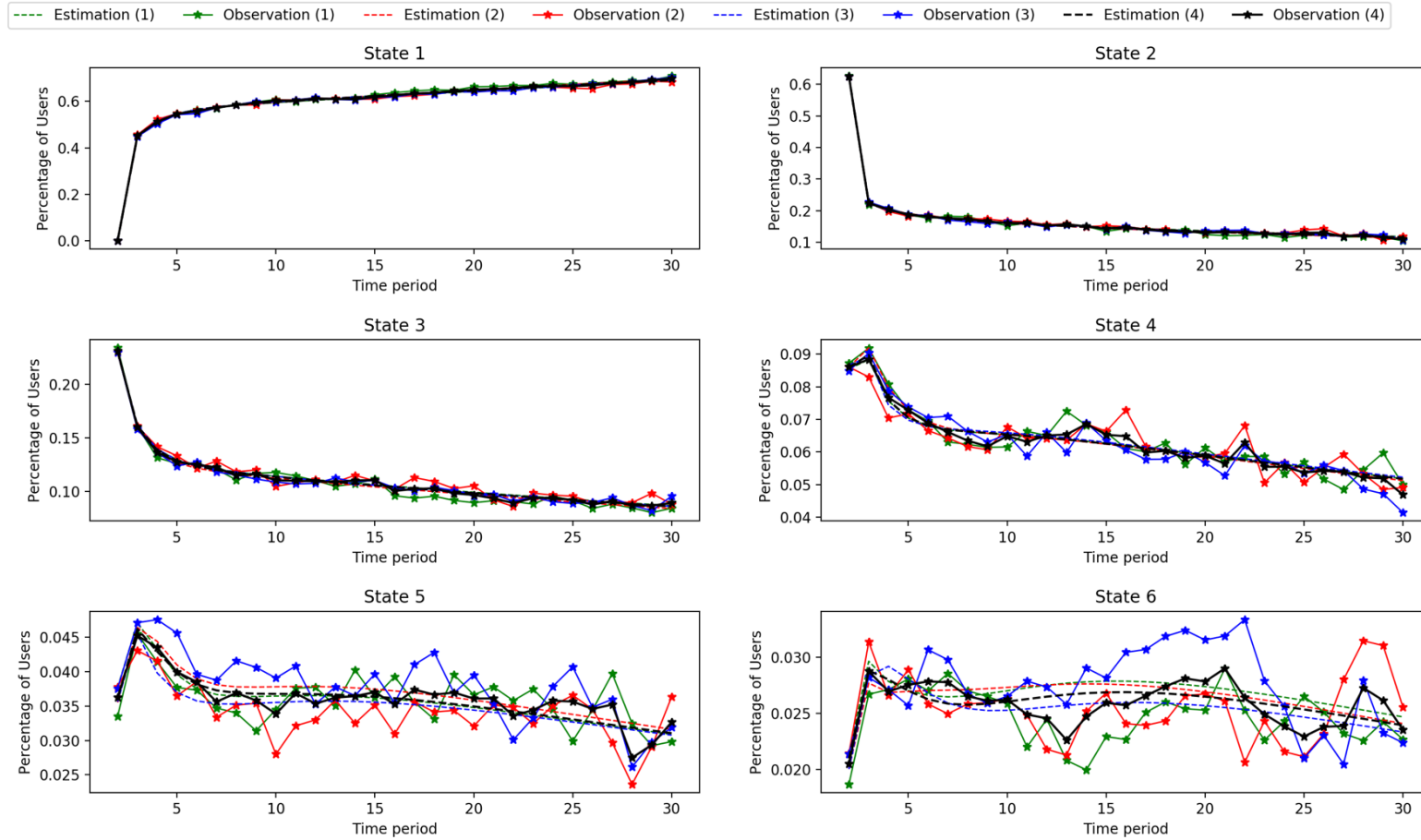


Figure 4 3-fold cross-validation in individual lifetime

Figure 4 illustrates the change in users over time in the respective states using different training and test data sets. We note that the y-axis in the six subfigures have different scales given the significant difference in the number of users. The lower frequency states have many more users than the higher frequency states. As the figures show, over time a majority of the users become inactive (enter State 1), even if correcting for departed users according to the methodology described in the appendix. For all active states (States 2-6) we observed downward trends. These were, however, least evident for the high frequency states (States 5 and 6), implying that there is a fairly constant population of high frequency users. The figures also show that fluctuations increase the higher the frequency state due to the reduction in user numbers. For States 1 to 4 (inactive up to medium usage), all three sets of data fit the estimation well. Overall, we conclude that the model appears to be reliable as we obtained similar low errors for all tests ($R(1) = 2.69\%$, $R(2) = 2.73\%$, $R(3) = 2.28\%$; $R(4) = 1.83\%$).

5.3 Estimation Results in Lifetime

The average values of the estimated parameters from the three tests are taken as estimation results and are shown in Table 5 (Stage Transition Matrix) and Table 6 (State transition matrices).

Table 5 Life Stage transition matrix π

Stage	B	G	M	R	D
B	-	0.670	-	0.330	-
G	-	0.639	0.361	-	-
M	-	-	0.987	0.013	-
R	-	-	-	0.908	0.092
D	-	-	-	-	1.000

$\pi_{BR} = 0.330$ indicates that 33% of users appeared to be dissatisfied with their carsharing trial and did not stay in the system. The remaining 67% of users develop some regular usage patterns. Further, in general, users tend to stay in the same stage, as we obtained $\pi_{GG}=0.639$, $\pi_{MM}=0.987$, $\pi_{RR}=0.908$. As expected, π_{MM} had the largest value, close to one. The value indicates that 98.7% of users in the maturity stage remain

from one month to the next in this stage. The π_{GG} estimate implies that that after four months nearly 80% of the users have already established their usage patterns and transferred to maturity.

We now turn to state transitions in Table 6. For illustrative purposes we apply colour coding with red to green indicating high to low values.

Table 6 State transition matrices for 5-stage model

Stage Birth						
State	1	2	3	4	5	6
1	-	0.626	0.231	0.086	0.036	0.021
Stage Growing						
State	1	2	3	4	5	6
1	0.618	0.255	0.108	0.018	0.001	0.000
2	0.475	0.322	0.162	0.039	0.002	0.000
3	0.200	0.392	0.296	0.112	0.000	0.000
4	0.110	0.147	0.205	0.330	0.174	0.034
5	0.036	0.036	0.083	0.185	0.361	0.299
6	0.017	0.049	0.058	0.034	0.146	0.696
Stage Maturity						
State	1	2	3	4	5	6
1	0.707	0.202	0.071	0.014	0.004	0.002
2	0.372	0.366	0.212	0.046	0.004	0.001
3	0.159	0.29	0.361	0.168	0.02	0.002
4	0.047	0.11	0.282	0.393	0.15	0.019
5	0.02	0.017	0.076	0.258	0.444	0.186
6	0.007	0.001	0.01	0.034	0.247	0.701
Stage Recession						
State	1	2	3	4	5	6
1	0.919	0.076	0.004	0.000	0.000	0.000
2	0.848	0.061	0.021	0.035	0.023	0.012
3	0.434	0.016	0.215	0.180	0.121	0.034
4	0.129	0.117	0.363	0.222	0.114	0.055
5	0.031	0.074	0.158	0.313	0.387	0.036
6	0.022	0.000	0.031	0.150	0.394	0.403

Firstly, looking at the state transition parameters in stage B, ϕ_B , we observed that it is much more likely that a user starts his usage with an initially low usage frequency. Only

5.6% used the scheme with a State 5 or 6 frequency in the initial month.

For Stage G, given the name ‘Growing’ we would expect a matrix biased towards larger values in the upper triangle, indicating that users increase their usage over time. Our results only partially confirmed this, however, as we observed that, for example, among those in State 4, 46.2% tended to decrease usage and only 20.8% tended to increase usage. We suggest this indicates that therefore this stage might be alternatively described as ‘Learning’ and that after an initial trial period with higher usage, many users then tend to settle into a fairly low-frequency state.

The diagonal values represent the probability of staying in the same state. Therefore constant usage would mean that in each the value on the diagonal equals 1; the larger the off-diagonal values, the more frequencies fluctuate. Accordingly, for life stage M, Maturity, we expect the diagonal values to dominate. We also observed that these values were indeed largest in each row with the exception of ϕ_{21}^M being slightly larger than ϕ_{22}^M . It is further noteworthy that the ϕ_{66}^M and the ϕ_{11}^M values are the largest. This indicates that a stable group of very high-frequency users exists as well as a group of temporarily inactive users. For States 2 to 5 the diagonal values are much lower, indicating that fluctuations to adjacent states are fairly frequent and that individuals adjust their behaviour gradually. ‘Transition jumps’ to non-adjacent states, i.e., from State 2 to 4, are instead much less frequent.

In Stage R we observe that values in the lower triangle of matrix Φ_R are indeed much larger. 85% of State 2 low-frequency users become inactive and once a user is inactive there is a 92% probability s/he remains inactive. Also most users will decrease usage once they enter the Recession Stage, i.e. ‘downwards jumps’ to States 2 or 1 are likely. However, we also observed that values of sudden transitions from low levels (2-3) to high levels (5-6) are larger. This may reflect a few users increasing usage briefly before ‘fully giving up’ and dropping out.

To support these findings we have added a comparison of state transitions for different stages in Table 7. Here red indicates larger positive values in the respective comparison, while blue represents the opposite. White indicates similar values in the two stages. We focus on the comparison of the three matrices for Stages G, M, R whereas we only have vectors for Stages B and D and furthermore the characteristics of these vectors are clear according to definitions. According to the ordering of the states, the transition matrix values below, on and above the diagonal represent the probability of transition to a lower, same and higher usage frequency state respectively. Therefore, the negative values on the diagonals in Table 7 indicate that the tendency for individuals in Stages G and R to remain at the same state is weaker than for those in the M stage. This supports our decision to refer to this stage as the ‘Maturity Stage’. Besides, positive values distributed in both the upper and lower triangle of Table 7a indicate that users in Stage G appear to be experimenting with different usage patterns, supporting our contention that this is an ‘Adaptation Stage’. In Table 7b instead few positive values were found in the upper triangle. This means that users were more likely to transfer to lower usage states Stage R compared with Stage M, bolstering the argument for referring to this stage as ‘Recession’.

Table 7 Stage comparison

<i>a) Difference (Stage G- Stage M)</i>						
State	1	2	3	4	5	6
1	-0.088	0.053	0.037	0.004	-0.003	-0.002
2	0.103	-0.044	-0.050	-0.007	-0.002	-0.001
3	0.041	0.103	-0.065	-0.056	-0.020	-0.002
4	0.063	0.037	-0.076	-0.063	0.024	0.015
5	0.016	0.019	0.008	-0.073	-0.083	0.113
6	0.010	0.048	0.049	0.000	-0.101	-0.005
<i>b) Difference (Stage R- Stage M)</i>						
State	1	2	3	4	5	6
1	0.212	-0.126	-0.067	-0.014	-0.004	-0.002
2	0.476	-0.305	-0.191	-0.011	0.019	0.011
3	0.275	-0.274	-0.146	0.012	0.101	0.032
4	0.082	0.007	0.081	-0.171	-0.036	0.036
5	0.011	0.057	0.082	0.055	-0.057	-0.150

6	0.015	-0.001	0.021	0.116	0.147	-0.298
---	-------	--------	-------	-------	-------	--------

Having established our methodology and estimation of transitions in user lifetime, in the following two sections we demonstrate two applications of the model. First we estimate transitions for different user groups and different time intervals. In the second application, we utilise the transitions to predict usage in system operation time.

5.4 Application 1: Influence of Previous Experience and Area Extension

With the development of the FFcs service, an increasingly larger service area has become available to service users in Montreal. The system extension in April 2015 was particularly significant, due to its city centre location, its large service area extension (to 78.1 km²) and the accompanying fleet extension (over 200 vehicles). This extension is likely to have an influence on the usage frequency as well and we therefore hypothesize that different models need to be estimated for the time before and after the extension.

Furthermore, it is important to note that the carsharing service operator first offered SBcs services before the FFcs service was added. Therefore some users of the FFcs service have prior experience of using the SBcs service. This experience may also affect user behaviour.

To analyse the influence of these two factors we denote the time from June 2013 to March 2015 as T1 and time periods thereafter as T2. Users who have been registered members of the SBcs before starting to use the FFcs service we denote as E (experienced users), and users without experience we denote as NE users. We then estimate four sets of model parameters for the user groups T1 E, T1 NE, T2 E, T2 NE. For the discussion we denote ΔTE as a comparison between T1 and T2 for E users, and similarly we refer to ΔTNE for a comparison of NE users before and after the extension. Furthermore, for a comparison of the effect of experience for the two different time periods we introduce $T1\Delta E$ and $T2\Delta E$. The estimated Stage Transition

Matrix is shown in Table 8. The longer tables with state transitions and differences in state transitions are shown in the appendix (see Table 9 and Table 10). Figure 5 shows the estimated state distribution curves by stages.

Table 8 Estimated parameters and differences in 4 conditions

T1 E	B	G	M	R	D	T2 E	B	G	M	R	D
B	-	0.563	-	0.437	-	B	-	0.599	-	0.401	-
G	-	0.721	0.279	-	-	G	-	0.636	0.364	-	-
M	-	-	0.998	0.002	-	M	-	-	0.999	0.001	-
R	-	-	-	0.955	0.045	R	-	-	-	0.933	0.067
D	-	-	-	-	1.000	D	-	-	-	-	1.000
T1 NE	B	G	M	R	D	T2 NE	B	G	M	R	D
B	-	0.681	-	0.319	-	B	-	0.829	-	0.171	-
G	-	0.394	0.606	-	-	G	-	0.247	0.753	-	-
M	-	-	0.952	0.048	-	M	-	-	0.976	0.024	-
R	-	-	-	0.965	0.035	R	-	-	-	0.831	0.169
D	-	-	-	-	1.000	D	-	-	-	-	1.000

We observe that in Stage B experienced users are more likely to briefly try the FFcs scheme (possibly out of curiosity) but then drop out and possibly remain with the SBcs scheme. This is evident from $\pi_{BR}^{T1,NE} > \pi_{BR}^{T1,E}$ and $\pi_{BR}^{T2,NE} > \pi_{BR}^{T2,E}$ indicating that more ‘jump transitions’ are occurring from Stage B to Stage R for non-experienced users both before and after the service extension. We suggest a possible explanation can be derived from the fact that the free-floating service has a different pricing structure: The free-floating service is attractive for short trips due to increased convenience in the form of minimized access time. However, this advantage is not significant for longer trips due to the higher price of the free-floating scheme. Therefore some users will quickly learn that the free-floating service is not attractive for their needs and decide to leave the service fairly soon, leading to a higher discontinuation ratio in early usage stages among experienced users.

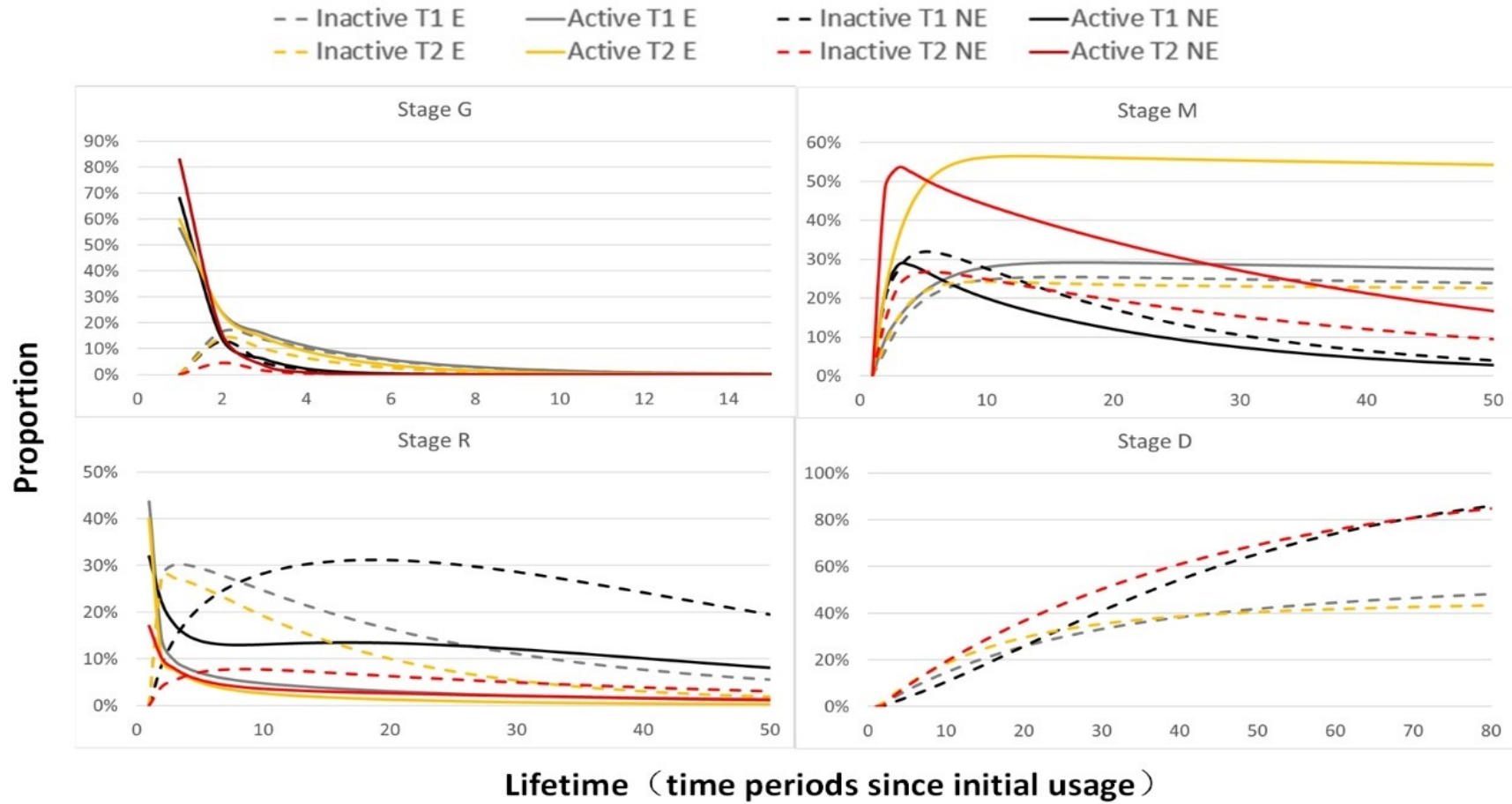


Figure 5 Stage Distributions in Lifetime distinguishing T1 and T2, NE and E as well as active and inactive users.

We further observe $\pi_{BR}^{T2,NE} < \pi_{BR}^{T1,NE}$ and $\pi_{BR}^{T2,E} < \pi_{BR}^{T1,E}$ meaning that, compared to T1, in T2, after the extension, fewer users enter the recession stage immediately, i.e., gain experience of the scheme and hence are more likely to become longer term users. Looking at the state transitions in more detail, we observed that non-experienced users were more likely to recorded higher usage in the initial time period (see Table 10 in the appendix, States 3-6 $T1\Delta E$ and $T2\Delta E$). Furthermore, we observed that immediate high frequency users increased in T2 for both E and NE users.

On average, E users stayed at stage G for 3.5 and 2.7 months, respectively, in T1 and T2, while NE stayed only 1.65 and 1.33 months in T1 and T2. This indicates that NE users display quicker ‘adaptation’, that is, formations of regular FFcs usage patterns. An explanation is that habits and satisfaction with the station-based scheme may deter users from quickly adapting to the new service. As both services meet some of their needs, the users do not perceive a need to increase usage of the free-floating service immediately, but the convenience of the free-floating service might lead to slower, gradual usage increases. It is further noteworthy that after the area extension, both user groups remained in stage G for shorter time periods. We suggest that these results are not unexpected.

According to estimated parameters $\pi_{MM}^{T1,NE}, \pi_{MM}^{T1,E}, \pi_{MM}^{T2,NE}, \pi_{MM}^{T2,E}$, E users could stay at Stage M for many years, while NE stayed on average only 21 months before the large extension (T1 NE) and 42 months after the large extension (T2 NE) in Stage M. Therefore, even though the experienced users have initially lower usage, longer adaption time and higher short-term departure ratios, from a long-term perspective these users of the station-based scheme are important for the operators of the free-floating service. Once they have remained through the initial phase, they are more likely to become loyal, long-term customers.

Furthermore, for both E and NE users, the service extension has not only led to increased overall usage but also had other positive impacts in that it led to longer durations of remaining in Stage M ($\pi_{MM}^{T2,NE} > \pi_{MM}^{T1,NE}, \pi_{MM}^{T2,E} > \pi_{MM}^{T1,E}$) as illustrated in Figure 5. We also note that in Stage M, referring to comparisons ΔTE and ΔTNE

shown in the appendix, most upper triangle values were positive, suggesting that after the facility extension stable, long-term users increase their overall usage frequency.

Analysis of Stage R helps to explain how long users remain (occasional) system members after starting to decrease usage. The expected recession periods for E users were 22 and 15 months at T1 and T2 respectively, whereas those for NE users were 28 and 6 months, i.e. NE users remained longer than E users in Stage R during T1 but relatively shorter than E users during R2. After the extension both groups remained for shorter periods in R, with the effect being more significant for NE users. Looking at the state transitions in this stage, we observe that NE users transferred more rapidly to low-frequency states. Taking these observations together, we conclude that there is a higher proportion of inactive users (not dropping out yet but no current usage) among the NE users.

Overall we conclude that NE users appear to be more eager to experience the service, they adapt more quickly and exhibit high-frequency usage at the beginning, but fewer are long-term high-frequency users. Instead experienced users show higher stability and loyalty. The service area extension has had various positive effects and attracted a higher proportion of short-time low-frequency users as well as long-term high-frequency users.

5.5 Application 2: Estimation and forecasting in System Operation Time

We next applied our models to estimate the total usage in system operation time. For this we only used the first 40 months for model estimation and forecast the remaining months. We also require an estimate of the number of new persons signing up to the scheme, that is $b(t)$, during the forecasted period as an input variable. Different estimation approaches are feasible based on the time series of new user observations until the point for which data are available. Zhang et al (2020) applied an estimation approach to new users that further took into account the dissemination of promotional information among the population and the interrelationship of new usage uptake with service availability. In the current study, because we have additional data on new users for the forecast period and because forecasting $b(t)$ is not our focus, we used this data

as input. We note that if $b(t)$ were to be estimated clearly the model fit in subsequent figures would be reduced.

The following figures show the results using the parameters from the ‘base model’ (Table 5 and Table 6) as well as results with the user group/time period differentiated parameters (Table 7 and Table 8 in the appendix). We denote the former as Estimation A (EA) and the latter as EB. The results before Month 40 and from Months 41 to 47 are estimation and forecasting, respectively. We would expect EB to be superior in performance since more information is used for estimation, but we want to verify this is also true for forecasting. In system operation time the user states are primarily of interest and are the focus of the subsequent discussion. For example, we refer to EA5 to denote the estimated users in State 5 with base model parameters at the system time t . Estimations for active users and usage state are shown in Figure 6 and Figure 7 respectively. O refers to observed data.

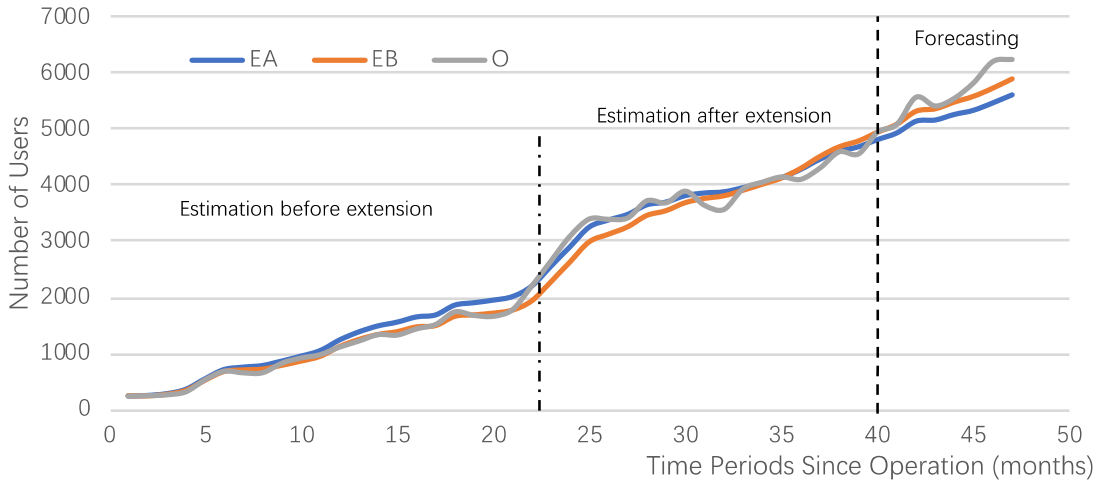


Figure 6 Observed and estimated/predicted active users

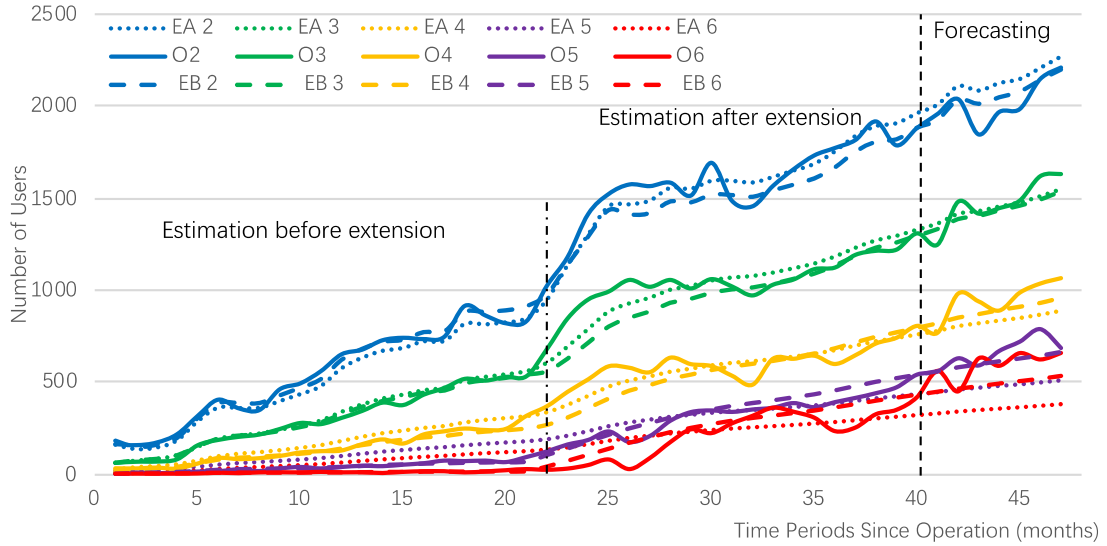


Figure 7 Estimation of active user states in system operation time

The RMSPE is again used to measure the performance of the two estimations and forecasting in system operation time. The errors for EA and EB are, respectively, 15.03% and 10.57%, which are both significantly higher than the values in the lifetime estimation. We find that, compared to EA, in EB the errors largely decrease in months 1-20 for estimation (from 20.1% to 9.2%), months 28-40 for estimation (from 10.9% to 8.2%), and months 41-47 for forecasting (from 14.1% to 7.1%), especially for high-frequency States 5 and 6. This shows that the prediction accuracy significantly increases when information about experience and area extension is utilized.

However, relatively large errors still exist in months 21 to 27 (EA: 15.5% EB: 19.1%). Analysing the data more closely, we find that this might be partially caused by the ‘revival’ of inactive users, as some users, including some without usage for several months, start to reuse the service again. This effect of ‘user resurrection’ is not considered in our model as D is an absorbing stage, that once entered cannot be left. We note that most of these users remain in low frequency states and are inactive again after a few months. We believe that due to this we observe more low and medium frequency users (State 2-4) than estimated during months 21 to 27. What is noteworthy is that the increase of users starts in Month 21, which is earlier than the area service extension in Month 22. Since the service provider releases information about upcoming extensions 1-2 months beforehand, we suggest that public communication

campaigns will stimulate (occasional) activity of the service even before the actual extension occurs.

6. CONCLUSIONS AND FURTHER WORK

This paper proposes an approach to describe the gradual change of behaviour over time with estimated transition probabilities between latent stages. The solution approach is based on Markovian updating. For estimating parameters, we first identify the states according to observation. We utilize the Expectation-Maximization (EM) algorithm to maximize the likelihood of fitting our observations with latent stages in the individual lifetime.

In Markovian approaches, the assumption of time-homogeneity, 1-step memory and a constant population are typically regarded as essential. However, time-homogeneity and short memory assumptions are not always suitable. Furthermore, the population of a system is dynamic due to new adopters. To address the problematic short-memory assumption, we draw on the idea of ‘life cycle stages’ to reflect the dynamically changing total number of potential users and we assume that the 1-step memory assumption only holds within stages. Accordingly, we propose a model with 5 stages. To address the dynamic population issue for system usage forecasting we discuss prediction in system operation time with new users per month as input. Finally, to partially address potential time-homogeneity issues we build separate models for different operational periods, in our case before and after service extension.

The major difference in the modelling approach compared to Zhang and Schmöcker (2017) is the presence of latent stages. We demonstrate that the model outperforms the model without latent stages, even controlling for the additional number of parameters. We show that this approach provides robust estimates of the usage patterns of carsharing users and suggest that this approach is useful for describing adaptation to a range of new transport solutions and significant infrastructure investments where users over time learn to value the new or improved system and start using it more frequently.

The model is applied to panel data from Montreal’s free-floating carsharing system. We

describe how users gradually change their usage frequency in early life stages, but how sharp reductions in usage can be seen before a user drops fully out. We observe that the semantics adopted for this paradigm of life stages fit the observations in general, except that the ‘Growth’ stage might be better termed as the ‘Learning’ phase, as users during this phase do not necessarily continuously increase usage. It is possible that users’ curiosity and desire to understand how best to utilize the free-floating scheme leads to higher usage in the beginning than is maintained over longer time periods.

We then apply the model to show the influence of previous experience and service extensions on user behaviour changes. We compare the differences in estimated parameters applied to four conditions (before / after area extension, with / without experience). We note that we can observe some marked differences between users with and without prior carsharing experience and that the significant service extension in May 2015 had a marked impact on the user lifetime and usage frequency. These findings imply that for usage predictions not only the quality of the new service itself but also experience with similar or previous services is important. This case study suggests that on one hand, experience and habitual use of the prior existing station-based service can serve as an ‘obstacle’ to adapt to the new, free-floating service, but it may also imply that experienced car-sharing users are in less need of the additional service since the station-based service already fulfills their needs. This manifests itself in a high discontinuation ratio in the first months after sign-up, initially a low usage frequency and a long adaptation time for those with station-based car sharing experience. On the other hand, once accepting the new service, these same users have higher usage frequency and longer lifetimes compared with those that use the free-floating service without prior station-based service experience. The practical implication for operators of a free-floating service is that, for long-term stable usage, attracting experienced users from a competing service should be the focus, even though for short-term benefit it is better to focus advertising on non-experienced users.

Besides such insights into expected ‘customer churn’ we use the estimation result and the observed number of new adopters to forecast the number of users in different states of the system operation time. We show that consideration of past experience and the service area extension decreases errors from 15.1% to 10.5% in estimation and 14.1%

to 7.1% in forecasting usage frequency states. We discuss that for further improvement research is required on how users in Stages R and D react to the service extension. We also acknowledge that our model is mainly intended for short-to medium-term predictions of an existing scheme for which data from some initial months are available. Understanding transferability of the estimates from this study to other contexts would require a range of additional case studies.

Methodologically the work presented here could be extended by introducing specific covariates for specific system internal or external events. For example, in the model without latent stages of Zhang and Schmöcker (2017) the month was considered as a covariate since the focus there was on showing shared bicycle usage fluctuations depending on seasonality. To gain additional insights into behavioural adjustments the model might be further extended by replacing user states based on frequency only with states based on ‘multidimensional factors’ including, besides frequency, spatial spread in usage and spread by time-of-day. We also consider extending this work by establishing the relationship between the demographic and geographic information and the parameters of the model. This will facilitate using the model for prediction of new transport schemes without historical data.

Acknowledgments

The authors would like to thank Communauto for providing the data needed for this research. We acknowledge the financial support of the Japan Society for the Promotion of Science (Project 18K04390). We further would like to thank Jane Singer for providing us with helpful editorial comments.

Conflict of Interest

All authors confirm that there is no conflict in interest.

REFERENCES

- Balać, M., Ciari, F., and Axhausen, K. W. (2015). Carsharing demand estimation: Zurich, Switzerland, area case study. *Transportation Research Record*, 2356, 10--18.
- Bartholomew, D.J. (1982). *Stochastic models for social processes* (3rd ed). London: Wiley.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer Science and Business Media.
- Burkhardt, J.E. and Millard-Ball, A. (2006). Who is attracted to carsharing? *Transportation Research Record*, 1986(1), 98--105.
- Catalano, M., Lo Casto, B. and Migliore, M. (2008). Car sharing demand estimation and urban transport demand modelling using stated preference techniques. *European Transport*, 40, 33-50.
- Cervero, R., and Tsai, Y. (2004). City CarShare in San Francisco, California: second-year travel demand and car ownership impacts. *Transportation Research Record*, 1887(1), 117--127.
- Ciari, F., and Axhausen, K.W. (2013). Carsharing Stated Preference (SP). *Travel Survey Metadata Series*, 47.
- Ciari, F., Balac, M. and Balmer, M. (2015). Modelling the effect of different pricing schemes on free-floating carsharing travel demand: a test case for Zurich, Switzerland. *Transportation*, 42(3), 413--433.

- Clewlrow, R.R. (2016). Carsharing and sustainable travel behaviour: Results from the San Francisco Bay Area. *Transport Policy*, 51, 158--164.
- Coll, M. H., Vandersmissen, M. H., and Thériault, M. (2014). Modelling spatio-temporal diffusion of carsharing membership in Québec City. *Journal of Transport Geography*, 38, 22--37.
- Costain, C., Ardron, C and Habib, K N. (2012). Synopsis of users' behaviour of a carsharing program: A case study in Toronto. *Transportation Research Part A: Policy and Practice*, 46(3), 421--434.
- Cox, W E. (1967). Product life cycles as marketing models. *The Journal of Business*, 40(4), 375--384.
- Dellaert, F. (2002). The expectation maximization algorithm. Georgia Institute of Technology. Available from < <https://www.cc.gatech.edu/~dellaert/em-paper.pdf>>.
- Efthymiou, D., Antoniou, C. and Waddell, P. (2013). Factors affecting the adoption of vehicle sharing systems by young drivers. *Transport policy*, 29, 64--73.
- Han, G., and Sohn, K. (2016). Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model. *Transportation Research Part B: Methodological*, 83, 121--135.
- Heilig, M., Mallig, N., Schröder, O., Kagerbauer, M. and Vortisch, P. (2015). Multiple-day agent-based modelling approach of station-based and free-floating carsharing. *Transportation Research Board 94th Annual Meeting*.
- Kalbfleisch, J.D and Lawless, J F. (1985). The analysis of panel data under a Markov assumption. *Journal of the American Statistical Association*, 80(392), 863--871.
- Kumar, V. and Reinartz, W. (2018). *Customer Relationship Management. Concepts, Strategy, and Tools*. Springer Texts in Business and Economics. Third Edition. ISBN 978-3-662-55380-0.
- Liu, F., Janssens, D., Cui, J., Wets, G. and Cools, M. (2015). Characterizing activity sequences using profile hidden Markov models. *Expert Systems with Applications*, 42(13), 5705--5722.
- Martin, E.W and Shaheen, S.A. (2011). Greenhouse gas emission impacts of carsharing in North America. *IEEE transactions on intelligent transportation systems*, 12(4), 1074--1086.
- Martinez, L.M., de Almeida Correia, G.H., Moura, F. and Lopes, M.M. (2017). Insights into carsharing demand dynamics: Outputs of an agent-based model application to Lisbon, Portugal. *International Journal of Sustainable Transportation*, 11(2), 148--

- Meijkamp, R. (1998). Changing consumer behaviour through eco-efficient services: an empirical study of car sharing in the Netherlands. *Business Strategy and the Environment*, 7(4), 234--244.
- Namaz, M., MacKenzie, D., Zerriffi, C. and Dowlatabadi, H. (2018). Is carsharing for everyone? Understanding the diffusion of carsharing services. *Transport Policy*, 63, 189--199.
- Rashwan, A.M., Kamel, M.S and Karray, F. (2013). Car driver fatigue monitoring using Hidden Markov Models and Bayesian networks. *Connected Vehicles and Expo (ICCVE), 2013 International Conference* (pp. 247--251). IEEE.
- Shaheen, S. and Cohen, A. (2016). Innovative Mobility Carsharing Outlook: Carsharing Market Overview, Analysis and Trends. Retrieved from Available from: "http://innovativemobility.org/wp-content/uploads/2016/02/Innovative-Mobility-Industry-Outlook_World-2016-Final.pdf". Accessed <November 2019>.
- Shaheen, S. A, Cohen, A.P and Chung, M.S. (2009). North American carsharing: 10-year retrospective. *Transportation Research Record*, 2110(1), 35--44.
- Shaheen, S.A and Schwartz, A. and Wiprywski, K. (2004). Policy considerations for carsharing and station cars: Monitoring growth, trends, and overall impacts. *Transportation Research Record*, 1887(1), 128--136.
- Shaheen, S.A, Sperling, D. and Wagner, C. (1999). Carsharing and partnership management: an international perspective. *Transportation Research Record*, 1666(1), 118--124.
- Singer, B. (1981). Estimation of nonstationary Markov chains from panel data. *Sociological Methodology*, 12, 319--337.
- Stathopoulos, A., Cirillo, C., Cherchi, E., Ben-Elia, E., Li, Y.-T. and Schmöcker, J.-D. (2017). Innovation adoption modelling in transportation: new models and data. 25, 61--68.
- Tang, K., Zhu, S., Xu, Y., and Wang, F. (2015). Analytical Modelling of Dynamic Decision-Making Behavior of Drivers During the Phase Transition Period Based on a Hidden Markov. *Transportation Research Board 94th Annual Meeting*. Washington D C.
- Watling, D. (1996). Asymmetric problems and stochastic process models of traffic assignment. *Transportation research. Part B: methodological*, 30(5), 339--357.

- Wielinski G., Trépanier, M. and Morency, C. (2019). Exploring service usage and activity space evolution in a free-floating carsharing service.. *Transportation Research Record*, Online.
- Yang, D., Xiong, C., Nasri, A. and Zhang, L. (2017). Hidden Markov Approach to Dynamically Modeling Car Ownership Behavior. *Transportation Research Record*, 2645(1), 123-130.
- Yoon, T., Cherry, C.R. and Jones, L.R. (2017). One-way and round-trip carsharing: A stated preference experiment in Beijing. *Transportation research part D: transport and environment*, 102--114.
- Yoon, T., Cherry, C.R., Ryerson, M.S., and Bell, J.E. (2019). Carsharing demand estimation and fleet simulation with EV adoption. *Journal of Cleaner Production*, 206, 1051--1058.
- Xiong, C., Chen, X., He, X. et al. (2015). The analysis of dynamic travel mode choice: a heterogeneous hidden Markov approach. *Transportation* 42, 985–1002.
- Xiong, C., Yang, D., and Zhang, L. (2018). A High-Order Hidden Markov Model and Its Applications for Dynamic Car Ownership Analysis. *Transportation Science*, 52(6), 1365--1375.
- Zhang, C. and Schmöcker, J.D. (2017). A Markovian model of user adaptation with case study of a shared bicycle scheme. *Transportmetrica B: Transport Dynamics*, 1--14.
- Zhang, C (2019). *Modelling the Demand Evolution of New Shared Mobility*. Ph.D. thesis, Kyoto University, Kyoto.
- Zhang, C., Schmöcker, J.D., Kuwahara, M., Nakamura, T., Uno, N (2020). A diffusion model to estimate adoption patterns of a one-way carsharing system in its initial years. *Transportation Research Part A: Policy and Practice*. 136, 135--150.
- Zhu, G., Song, K., Zhang, P. and Wang, Li. (2016). A traffic flow state transition model for urban road network based on Hidden Markov Model. *Neurocomputing*, 214, 567--574.

Appendix: Detailed Estimation Results for Application 2

Table 9 Estimated parameters for state transitions

<i>Stage Birth</i>													
T1 E	1	2	3	4	5	6	T2 E	1	2	3	4	5	6
1	-	0.724	0.207	0.057	0.010	0.002	1	-	0.718	0.197	0.059	0.020	0.005
T1 NE	1	2	3	4	5	6	T2 NE	1	2	3	4	5	6
1	-	0.548	0.291	0.115	0.037	0.009	1	-	0.455	0.276	0.133	0.077	0.058
<i>Stage Growing</i>													
T1 E	1	2	3	4	5	6	T2 E	1	2	3	4	5	6
1	0.660	0.248	0.078	0.013	0.000	0.000	1	0.642	0.282	0.065	0.006	0.001	0.003
2	0.467	0.327	0.158	0.041	0.006	0.002	2	0.495	0.308	0.134	0.044	0.016	0.004
3	0.248	0.289	0.299	0.135	0.025	0.004	3	0.174	0.354	0.305	0.132	0.032	0.003
4	0.094	0.18	0.277	0.327	0.122	0.000	4	0.037	0.166	0.299	0.343	0.130	0.026
5	0.015	0.113	0.204	0.377	0.292	0.000	5	0.026	0.068	0.103	0.294	0.365	0.143
6	0.003	0.018	0.077	0.215	0.507	0.181	6	0.016	0.107	0.045	0.007	0.087	0.739
T1 NE	1	2	3	4	5	6	T2 NE	1	2	3	4	5	6
1	0.449	0.37	0.148	0.023	0.007	0.004	1	0.684	0.208	0.098	0.009	0.000	0.000
2	0.451	0.302	0.189	0.052	0.006	0.000	2	0.513	0.269	0.145	0.054	0.011	0.008
3	0.179	0.376	0.259	0.163	0.023	0.000	3	0.204	0.252	0.320	0.157	0.052	0.015
4	0.077	0.186	0.326	0.347	0.064	0.000	4	0.083	0.112	0.252	0.305	0.185	0.063
5	0.031	0.063	0.187	0.114	0.529	0.075	5	0.039	0.049	0.106	0.203	0.386	0.217
6	0.022	0.000	0.000	0.234	0.000	0.744	6	0.026	0.013	0.036	0.034	0.239	0.652
<i>Stage Maturity</i>													
T1 E	1	2	3	4	5	6	T2 E	1	2	3	4	5	6
1	0.637	0.267	0.086	0.009	0.000	0.000	1	0.695	0.226	0.06	0.011	0.005	0.003
2	0.376	0.361	0.208	0.055	0.000	0.000	2	0.387	0.354	0.212	0.045	0.000	0.002
3	0.174	0.287	0.349	0.178	0.011	0.000	3	0.157	0.284	0.372	0.165	0.019	0.004
4	0.076	0.110	0.328	0.401	0.074	0.011	4	0.031	0.107	0.275	0.404	0.158	0.025
5	0.002	0.006	0.118	0.371	0.419	0.084	5	0.016	0.003	0.05	0.231	0.499	0.201
6	0.001	0.066	0.058	0.036	0.342	0.498	6	0.006	0.000	0.000	0.06	0.293	0.640
T1 NE	1	2	3	4	5	6	T2 NE	1	2	3	4	5	6
1	0.654	0.259	0.078	0.008	0.000	0.000	1	0.725	0.186	0.063	0.015	0.007	0.004
2	0.416	0.344	0.198	0.040	0.002	0.000	2	0.373	0.354	0.209	0.052	0.009	0.003
3	0.211	0.307	0.314	0.151	0.016	0.000	3	0.139	0.262	0.367	0.191	0.034	0.007
4	0.052	0.147	0.328	0.349	0.124	0.000	4	0.045	0.082	0.28	0.398	0.167	0.028
5	0.021	0.057	0.207	0.293	0.355	0.067	5	0.021	0.025	0.069	0.228	0.462	0.195
6	0.006	0.003	0.112	0.000	0.367	0.511	6	0.007	0.000	0.003	0.035	0.232	0.722

<i>Stage Recession</i>													
T1 E	1	2	3	4	5	6	T2 E	1	2	3	4	5	6
1	0.035	0.663	0.207	0.049	0.045	0.000	1	0.368	0.052	0.375	0.149	0.041	0.015
2	0.991	0.000	0.009	0.000	0.001	0.000	2	1.000	0.000	0.000	0.000	0.000	0.000
3	0.015	0.655	0.173	0.152	0.002	0.003	3	0.128	0.000	0.334	0.327	0.134	0.077
4	0.005	0.03	0.589	0.031	0.229	0.115	4	0.031	0.000	0.497	0.287	0.161	0.024
5	0.001	0.000	0.022	0.072	0.402	0.504	5	0.007	0.000	0.861	0.121	0.000	0.012
6	0.001	0.001	0.128	0.831	0.001	0.039	6	0.002	0.000	0.912	0.000	0.086	0.001
T1 NE	1	2	3	4	5	6	T2 NE	1	2	3	4	5	6
1	0.904	0.059	0.016	0.021	0.000	0.000	1	1	0.000	0.000	0.000	0.000	0.000
2	1.000	0.000	0.000	0.000	0.000	0.000	2	0.216	0.391	0.381	0.012	0	0.000
3	0.664	0.000	0.285	0.000	0.048	0.003	3	0.754	0.246	0.000	0.000	0.000	0.000
4	0.46	0.000	0.000	0.08	0.314	0.146	4	0.339	0.616	0.000	0.000	0.045	0.000
5	0.02	0.000	0.000	0.696	0.006	0.278	5	0.044	0.09	0.136	0.396	0.334	0.000
6	0.02	0.049	0.000	0.293	0.627	0.01	6	0.039	0.06	0.228	0.204	0.14	0.329

Table 10 Difference between estimated parameters in 4 conditions
(Red (blue) indicates larger (smaller) values in the respective comparison, white similar values.)

Stage Birth													
T1ΔE	1	2	3	4	5	6	T2ΔE	1	2	3	4	5	6
1	-	-0.176	0.084	0.058	0.027	0.007	1	-	-0.263	0.079	0.074	0.057	0.053
ΔTE	1	2	3	4	5	6	ΔTNE	1	2	3	4	5	6
1	-	-0.006	-0.010	0.002	0.010	0.003	1	-	-0.093	-0.015	0.018	0.04	0.049
Stage Growing													
T1ΔE	1	2	3	4	5	6	T2ΔE	1	2	3	4	5	6
1	-0.211	0.122	0.070	0.010	0.007	0.004	1	0.042	-0.074	0.033	0.003	-0.001	-0.003
2	-0.016	-0.025	0.031	0.011	0.000	-0.002	2	0.018	-0.039	0.011	0.010	-0.005	0.004
3	-0.069	0.087	-0.040	0.028	-0.002	-0.004	3	0.030	-0.102	0.015	0.025	0.020	0.012
4	-0.017	0.006	0.049	0.020	-0.058	0.000	4	0.046	-0.054	-0.047	-0.038	0.055	0.037
5	0.016	-0.050	-0.017	-0.263	0.237	0.075	5	0.013	-0.019	0.003	-0.091	0.021	0.074
6	0.019	-0.018	-0.077	0.019	-0.507	0.563	6	0.010	-0.094	-0.009	0.027	0.152	-0.087
ΔTE	1	2	3	4	5	6	ΔTNE	1	2	3	4	5	6
1	0.235	-0.162	-0.05	-0.014	-0.007	-0.004	1	-0.018	0.034	-0.013	-0.007	0.001	0.003
2	0.062	-0.033	-0.044	0.002	0.005	0.008	2	0.028	-0.019	-0.024	0.003	0.010	0.002
3	0.025	-0.124	0.061	-0.006	0.029	0.015	3	-0.074	0.065	0.006	-0.003	0.007	-0.001
4	0.006	-0.074	-0.074	-0.042	0.121	0.063	4	-0.057	-0.014	0.022	0.016	0.008	0.026
5	0.008	-0.014	-0.081	0.089	-0.143	0.142	5	0.011	-0.045	-0.101	-0.083	0.073	0.143
6	0.004	0.013	0.036	-0.2	0.239	-0.092	6	0.013	0.089	-0.032	-0.208	-0.420	0.558
Stage Maturity													
T1ΔE	1	2	3	4	5	6	T2ΔE	1	2	3	4	5	6
1	0.017	-0.008	-0.008	-0.001	0.000	0.000	1	0.030	-0.040	0.003	0.004	0.002	0.001
2	0.040	-0.017	-0.010	-0.015	0.002	0.000	2	-0.014	0.000	-0.003	0.007	0.009	0.001
3	0.037	0.020	-0.035	-0.027	0.005	0.000	3	-0.018	-0.022	-0.005	0.026	0.015	0.003
4	-0.024	0.037	0.000	-0.052	0.050	-0.011	4	0.014	-0.025	0.005	-0.006	0.009	0.003
5	0.019	0.051	0.089	-0.078	-0.064	-0.017	5	0.005	0.022	0.019	-0.003	-0.037	-0.006
6	0.005	-0.063	0.054	-0.036	0.025	0.013	6	0.001	0.000	0.003	-0.025	-0.061	0.082
ΔTE	1	2	3	4	5	6	ΔTNE	1	2	3	4	5	6
1	0.071	-0.073	-0.015	0.007	0.007	0.004	1	0.058	-0.041	-0.026	0.002	0.005	0.003
2	-0.043	0.010	0.011	0.012	0.007	0.003	2	0.011	-0.007	0.004	-0.010	0.000	0.002
3	-0.072	-0.045	0.053	0.040	0.018	0.007	3	-0.017	-0.003	0.023	-0.013	0.008	0.004
4	-0.007	-0.065	-0.048	0.049	0.043	0.028	4	-0.045	-0.003	-0.053	0.003	0.084	0.014
5	0.000	-0.032	-0.138	-0.065	0.107	0.128	5	0.014	-0.003	-0.068	-0.140	0.080	0.117
6	0.001	-0.003	-0.109	0.035	-0.135	0.211	6	0.005	-0.066	-0.058	0.024	-0.049	0.142
Stage Recession													
T1ΔE	1	2	3	4	5	6	T2ΔE	1	2	3	4	5	6
1	0.869	-0.604	-0.191	-0.028	-0.045	0.000	1	0.632	-0.052	-0.375	-0.149	-0.041	-0.015
2	0.009	0.000	-0.009	0.000	-0.001	0.000	2	-0.784	0.391	0.381	0.012	0.000	0.000
3	0.649	-0.655	0.112	-0.152	0.046	0.000	3	0.626	0.246	-0.334	-0.327	-0.134	-0.077
4	0.455	-0.030	-0.589	0.049	0.085	0.031	4	0.308	0.616	-0.497	-0.287	-0.116	-0.024
5	0.019	0.000	-0.022	0.624	-0.396	-0.226	5	0.037	0.090	-0.725	0.275	0.334	-0.012
6	0.019	0.048	-0.128	-0.538	0.626	-0.029	6	0.037	0.060	-0.684	0.204	0.054	0.328
ΔTE	1	2	3	4	5	6	ΔTNE	1	2	3	4	5	6
1	0.096	-0.059	-0.016	-0.021	0.000	0.000	1	0.333	-0.611	0.168	0.100	-0.004	0.015
2	-0.784	0.391	0.381	0.012	0.000	0.000	2	0.009	0.000	-0.009	0.000	-0.001	0.000
3	0.090	0.246	-0.285	0.000	-0.048	-0.003	3	0.113	-0.655	0.161	0.175	0.132	0.074
4	-0.121	0.616	0.000	-0.080	-0.269	-0.146	4	0.026	-0.030	-0.092	0.256	-0.068	-0.091
5	0.024	0.090	0.136	-0.300	0.328	-0.278	5	0.006	0.000	0.839	0.049	-0.402	-0.492
6	0.019	0.011	0.228	-0.089	-0.487	0.319	6	0.001	-0.001	0.784	-0.831	0.085	-0.038