IIE Transactions

ISSN: 0740-817X (Print) 1545-8830 (Online) Journal homepage: http://www.tandfonline.com/loi/uiie20

Bike Sharing Systems: User Dissatisfaction in the Presence of Unusable Bicycles

Mor Kaspi, Tal Raviv & Michal Tzur

To cite this article: Mor Kaspi, Tal Raviv & Michal Tzur (2016): Bike Sharing Systems: User Dissatisfaction in the Presence of Unusable Bicycles, IIE Transactions, DOI: 10.1080/0740817X.2016.1224960

To link to this article: http://dx.doi.org/10.1080/0740817X.2016.1224960

Accepted author version posted online: 23 Aug 2016.
Published online: 23 Aug 2016.

Bike Sharing Systems: User Dissatisfaction in the Presence of Unusable Bicycles

Mor Kaspi, Tal Raviv, Michal Tzur

Department of Industrial Engineering, Faculty of Engineering, Tel-Aviv University, Israel

September 2015, revised: January 2016, March 2016, July 2016

Abstract

In bike sharing systems, at any given moment, a certain share of the bicycle fleet is unusable. This phenomenon may highly affect the quality of service provided to the users. However, this matter has not received so far any attention in the literature. In this study, the users' quality of service is modeled by their satisfaction from the system. We measure user dissatisfaction by a weighted sum of the expected shortages of bicycles and lockers in a single station. The shortages are evaluated as a function of the initial inventory of usable and unusable bicycles in the station. We analyze the convexity of the resulting bivariate function and propose an accurate method for fitting a convex polyhedral function to it. The fitted polyhedral function can later be used in linear optimization models for operational and strategic decision making in bike sharing systems. Our numerical results demonstrate the significant effect of the presence of unusable bicycles on user dissatisfaction. This emphasizes the need for having accurate real-time information regarding bicycle usability.

Key words

Bike sharing, Maintenance, Discrete convex analysis, Discrete optimization

1. Introduction

Bike sharing systems are nowadays operating in more than 900 cities around the world (DeMaio and Meddin, 2015). One essential key to the sustainability of a bike sharing system is efficient planning of maintenance operations. Along with their fast implementation, bike sharing systems are receiving a growing attention in the operations management and operations research literature. However, to our knowledge, maintenance aspects have not been studied yet.

Previous studies dealt with a range of topics varying from aspects regarding the design of the system to operational issues. The design of the system includes determining the location of the stations, the capacity (number of lockers, also known as docking points) of each station and the fleet size, see for example Lin and Yang (2011), Lin et al. (2013), George and Xia (2011), and Shu et al. (2013).

The difficulty in operating bike-sharing systems arises from the need to constantly satisfy the demand for bicycles and for lockers. The demand processes are typically asymmetric and time heterogeneous. Due to this nature, shortages of bicycles or lockers might occur in some stations along the day. In order to provide good service to the users, the operators strive to reduce the occurrences of shortages. The most studied topic so far in the bike sharing literature concerns the rebalancing of bicycle inventory levels in the stations. This topic can be divided into two related operational planning decisions. The first is concerned with determining the target inventory level in each station at the beginning of the day, see Raviv and Kolka (2013), Schuijbroek et al. (2013) and Vogel et al. (2014). The second is concerned with planning the routing of repositioning trucks between the stations and the loading/unloading operations. This can be done either during the night when the system is nearly idle and the traffic is low, with the goal of preparing the system for the next working day (static repositioning), or during the day when it is possible to react to unexpected events (dynamic repositioning). For studies on static repositioning see, for example, Nair and Miller-Hooks (2011), Benchimol et al. (2011), Angeloudis et al. (2012), Chemla et al. (2013a), Raviv et al. (2013), Dell'Amico et al. (2014), Erdoğan et al. (2014), Erdoğan et al. (2015) and Forma et al. (2015). Few studies focus on dynamic repositioning, see Contardo et al. (2012), Kloimüllner et al. (2014) and Pessach et al. (2014).

Another approach is to ease the imbalance problem by implementing system regulations or policies. Fricker and Gast (2014) propose a best-of-two regulation under which a user who returns a vehicle is directed to the least congested station between two preferred ones. Kaspi et al. (2014, 2016c) suggest implementing parking reservation policies as means of reducing users' dissatisfaction and uncertainty by redirecting them to less congested stations. Chemla et al. (2013b) and Pfrommer et al. (2014) study pricing mechanisms.

An underlying assumption in all of the above studies is that all bicycles in the system are at a usable state at all times. In practice, however, every day a certain number of bicycles become unusable and require repair. These bicycles block some of the system resources, i.e., the lockers. Therefore, it is important to monitor and report the usability level of bicycles. The percentage of bicycles that become unusable every day is certainly not negligible and therefore this phenomenon should be taken into consideration in the planning of daily operational activities. For example, in 2009, an independent inspection carried in stations of Bixi Montreal, has shown that one out of five bicycles was in a disrepair state (DeMaio and Meddin 2009). In 2009, a study of the city of Barcelona on the Bicing bike sharing system reported that 12% of all available bicycles had major defects that prevented normal usage and 55% had minor defects such as broken bells and lights (Castro, 2011). In 2014, the Hangzhou bike-sharing system had 78,000 bicycles in its fleet and nearly 2,000 bicycles required maintenance every day (DeMaio and Meddin, 2014).

Unfortunately, most bike-sharing systems operators do not publish (or even do not collect) data regarding unusable bicycles. To the best of our knowledge, the only bike sharing operator that regularly publishes maintenance reports is NYC Bikeshare, the operator of Citibike, the bike sharing system in New-York: https://www.citibikenyc.com/system-data/operating-reports. During 2014, nearly 330 stations were operating in the system and the average fleet size was 5,137 bicycles. In total, 8,791,987 trips were executed, 528,319 bicycles were repositioned and 34,806 bicycles were repaired in two designated workshops. That is, on average, 95 bicycles needed major repairing every day, which stands for almost 2% of the fleet. These figures exclude many bicycles that were repaired in field (the exact numbers of in field repairs are not reported). Each bicycle that requires major repairing is repositioned at least twice (to and from the workshop), that is, at least 13% of the repositioning efforts are due to maintenance operations. We suspect that in many other systems, especially older ones or those who do not monitor or publish repair records, the maintenance operations may require even more resources.

Unusable bicycles have several implications: (a) until they are removed by repositioning workers or are fixed at the stations, they block lockers and therefore reduce the number of usable lockers in the stations. (b) In some cases the unusable bicycles must be removed from the stations and moved to the workshop. These bicycles require repositioning resources that otherwise could be used for rebalancing the system. This should be taken into account when planning repositioning operations. (c) Incomplete or inaccurate information regarding the usability of bicycles may result with inferior decision making of both the operators and the users.

Detection of unusable bicycles (addressing (c)) and integration of the collection of these bicycles with repositioning activities (addressing (b)) are at the focus of two other studies we carried out in parallel. In Kaspi et al. (2016b) we develop a probabilistic model for the detection of unusable bicycles, which can be incorporated in the on-line information system. In Kaspi et al. (2016a) we propose an optimization model that integrates the collection of unusable bicycles with the repositioning activities, as well as a dynamic programming model that allows updating decisions at each station on the route, as the actual number of unusable bicycles is revealed.

The study described in the current paper examines the effect of unusable bicycles on the service level provided in a single station (addressing the implications of (a)). Obviously, the effect of the presence of unusable bicycles may differ between stations, depending on their capacities and the demand processes for bicycles and lockers. A better understanding of this effect will assist in better planning their collection. Clearly, in order to maximize the service level, all unusable bicycles should be removed from the stations or be repaired as soon as possible. However, since the transportation and maintenance resources of the operator are limited, a method to estimate the expected effect of the unusable bicycles at each station can help in prioritizing these operations.

The contribution of this paper is as follows: we introduce an *Extended User Dissatisfaction Function* (*EUDF*) that represents the expected weighted number of users that are unable to rent or return a bicycle during a given period as a bivariate function of the initial number of usable and unusable bicycles at the station. This is an extension of the *User Dissatisfaction Function*

(*UDF)* that was initially presented in Raviv and Kolka (2013), which assumed that bicycles are always usable. We prove some discrete convexity properties of the EUDF. In addition, we propose a method for calculating a convex polyhedral function that has nearly identical values as the EUDF in its range. This polyhedral approximation can be used to optimize the initial bicycle inventories in the system subject to various constraints using linear programming, see for example Kaspi (2015, pp. 105-108).

The remainder of this paper is structured as follows. In Section 2 a formulation of the EUDF for a single station is presented. Properties of the EUDF, including a convexity analysis are provided in Section 3. In Section 4, a method to approximate the EUDF by a convex polyhedral function is presented. Results of a numerical experiment that examines the accuracy of the approximation are reported in Section 5. Concluding remarks are given in Section 6.

2. Extended User Dissatisfaction Function

In this section we present an extension of the UDF that was introduced by Raviv and Kolka (2013). Assuming that a station is not visited by repositioning vehicles throughout a given period (say, a day), the UDF represents the user dissatisfaction (a measure for the service level) as a discrete function of the initial number of bicycles in the station. Specifically, the user dissatisfaction is expressed as a weighted sum of the expected shortages of bicycles and the expected shortages of lockers along the given period. In Raviv and Kolka (2013) it is proven that the UDF is a convex function of the initial inventory of bicycles.

We begin by describing some notation and modeling assumptions that were presented in Raviv and Kolka (2013), which are also a part of the extended model. Subsequently, we present some revised notation and additional assumptions needed for the extended model. A list of the notations presented here and henceforward is given in Appendix A1 (Table 7).

We model a single station during a finite period $[0,T]$. Initially, at time 0, there is a certain number of bicycles in the station. During this period, users who wish to rent or return a bicycle arrive at the station according to an arbitrary stochastic process. If the demand for a bicycle/available locker can be satisfied, the bicycle is rented/returned and the station's inventory level is updated accordingly. On the other hand, if the demand cannot be satisfied, we assume that the user immediately abandons the station (she either roams to a nearby station or abandons the system). We note that the model does not take into account mutual influences of neighboring stations on the demand process.

There are two sources for user dissatisfaction, namely shortage of bicycles and shortage of lockers. For each type of shortage, the system is penalized by an amount that represents the dissatisfaction caused due to this shortage. We denote by p the penalty for each user who faces shortage of bicycles and by h the penalty for each user who faces shortage of lockers. The total number of lockers in the station is denoted by C. We refer to this value as the capacity of the station.

In the EUDF an additional dimension is introduced. The initial inventory of bicycles is divided into two groups, namely, usable and unusable (broken) bicycles, denoted by I_0 and B_0, respectively. This extension allows us to examine the effect of changes in the initial inventory level of each group on the service level, given their joint station capacity. In particular, the effect of the presence of unusable bicycles can be studied. However, the analysis of the EUDF becomes more difficult, as will be described in Section 3.

We assume that during the given period, the inventory level is not externally altered, that is, until time T, no repositioning or repairing activities are performed in the station. In particular, this implies that the number of unusable bicycles in the station cannot decrease during the given period, since it is assumed that unusable bicycles would not be rented by the users. However, some bicycles may become unusable during the ride, that is, some bicycles may be returned unusable to the station. Therefore, the number of unusable bicycles in the station may increase during the given period. Lastly, we assume that there is no change in the condition of the bicycles while they are parked in the station. Indeed, in some cases, bicycles are damaged when they are parked in the station due to vandalism or weather conditions. However, these occurrences are very rare and thus are negligible. In some systems, the collection of unusable bicycles may be done separately from the repositioning activities. Therefore, the inventory level of unusable bicycles may decrease during the given period due to the collection activity. In Appendix A2, we demonstrate how the EUDF can be adapted to such scenarios.

Let E^R denote the time epochs in which the demands for bicycles or lockers occur under demand realization R. We denote by $I_j^R\left(I_0, B_0\right)$ the inventory level of usable bicycles right after the j^{th} demand occurrence under realization R, given the inventory of usable and unusable bicycles $\left(I_0, B_0\right)$ at time 0. For the sake of brevity, we omit the conditioning on the initial inventory in subsequent notation. Similarly, the inventory level of unusable bicycles right after the j^{th} demand occurrence under realization R is denoted by $B_j^R\left(I_0, B_0\right)$. The demand for bicycles or lockers at the j^{th} occurrence is denoted by the pair $\left(d_j^{R,I}, d_j^{R,B}\right) \in \left\{(1,0), (-1,0), (0,-1)\right\}$, where (1,0) represents a demand for a usable bicycle, (-1,0) represents a demand for a locker in order to return a usable bicycle and (0,-1) represents a demand for a locker in order to return an unusable bicycle. Next, we present a recursive function, denoting the number of usable bicycles in the station after the occurrence of the j^{th} demand, given the inventory levels after the $\left(j-1\right)^{st}$ demand occurrence:

$$I_j^R\left(I_0,B_0\right)=\begin{cases}0 & I_{j-1}^R\left(I_0,B_0\right)-d_j^{R,I}<0\\ C-B_{j-1}^R\left(I_0,B_0\right) & I_{j-1}^R\left(I_0,B_0\right)-d_j^{R,I}>C-B_{j-1}^R\left(I_0,B_0\right)\\ I_{j-1}^R\left(I_0,B_0\right)-d_j^{R,I} & otherwise\end{cases}$$

And the number of unusable bicycles in the station is given by the following:

$$B_j^R\left(I_0,B_0\right)=\begin{cases}C-I_{j-1}^R\left(I_0,B_0\right) & B_{j-1}^R\left(I_0,B_0\right)-d_j^{R,B}>C-I_{j-1}^R\left(I_0,B_0\right)\\ B_{j-1}^R\left(I_0,B_0\right)-d_j^{R,B} & otherwise\end{cases}$$

We refer to $C-B_j^R\left(I_0,B_0\right)$ as the *effective capacity* of the station. Since unusable bicycles cannot be rented, they block the lockers in which they are parked.

Let $\Delta_j^R\left(I_0,B_0\right)$ and $\Theta_j^R\left(I_0,B_0\right)$ be indicator functions that indicate whether a user faces shortage of a bicycle or a locker as the j^{th} demand occurs. Let $\left(x\right)^+=\max\left\{0,x\right\}$. Then, the bicycle shortage indicator is given by $\Delta_j^R\left(I_0,B_0\right)=\left(-I_{j-1}^R\left(I_0,B_0\right)+d_j^{R,I}\right)^+$ and the locker shortage indicator is given by $\Theta_j^R\left(I_0,B_0\right)=\left(I_{j-1}^R\left(I_0,B_0\right)+B_{j-1}^R\left(I_0,B_0\right)-d_j^{R,I}-d_j^{R,B}-C\right)^+$.

We denote by $F^R\left(I_0,B_0\right)$ the total dissatisfaction of users under demand realization R. The total dissatisfaction is obtained by summing all the shortages for bicycles and lockers and multiplying each shortage by the related penalty:

$$F^R(I_0,B_0)=\sum_{j=1}^{|E^R|}\left(p\cdot \Delta_j^R(I_0,B_0)+h\cdot \Theta_j^R(I_0,B_0)\right)$$

Then, we denote by $F(I_0,B_0)$ the *expected* penalty (over all realizations) due to shortages of bicycles and lockers during the given period as a discrete function of the initial inventory of usable (I_0) and unusable (B_0) bicycles. We refer to this function as the EUDF and it is given by the following equation:

$$F(I_0,B_0)\equiv \mathbb{E}_R\left\{F^R(I_0,B_0)\right\}=\mathbb{E}_R\left\{\sum_{j=1}^{|E^R|}\left(p\cdot \Delta_j^R(I_0,B_0)+h\cdot \Theta_j^R(I_0,B_0)\right)\right\} \quad (1)$$

Note that the UDF is a special case of this model in which $B_0=0$ and there is no demand for lockers in order to return unusable bicycles.

3. Analysis of the EUDF

In this section we analyze the EUDF (1) and study its convexity, which is helpful for optimization purposes. In Section 3.1 we prove several properties of the EUDF, which are later used in its convexity analysis, presented in Section 3.2.

3.1. Properties of the EUDF

We begin our analysis of the EUDF by proving that it is non-decreasing in the initial inventory of unusable bicycles. Intuitively, this is true because an addition of unusable bicycles decreases the effective capacity of the station. We next prove this observation formally.

We denote by Ω a sequence of demand occurrences which do not include a returning of an unusable bicycle. Note that Ω can represent an entire demand realization or a subset of it. Thus,

$F^{\Omega}(I_0, B_0)$ denotes the user dissatisfaction under this sequence of demand occurrences. After analyzing such sequences, we will extend our analysis to any demand realization.

Lemma 1: For any sequence of demand occurrences Ω, the following inequality holds: $F^{\Omega}(I_0, B_0+1) \geq F^{\Omega}(I_0, B_0)$.

Proof: Consider two initial settings of a station: (I_0, B_0+1) and (I_0, B_0), i.e., when the number of usable bicycles is identical, but the number of unusable bicycles differs by one. We claim that the number of usable bicycles under both settings may differ by at most 1 at any time, namely, either $I_j^{\Omega}(I_0, B_0+1) = I_j^{\Omega}(I_0, B_0)$ or $I_j^{\Omega}(I_0, B_0+1) = I_j^{\Omega}(I_0, B_0) - 1$. This will be demonstrated using Table 1. In Table 1 we describe four different shortage events that may occur in either of these settings. In the second column we present the relations between usable inventory levels after the $(j-1)^{st}$ demand occurrence under settings (I_0, B_0) and (I_0, B_0+1). In the third column we describe the type of shortage, namely, bicycle or locker. In the fourth and fifth columns we denote under which setting this shortage occurs. Finally, in the sixth column we present the relations between the inventory levels under the two settings that result from the shortage event. Note that at time 0, the inventory levels of usable bicycles are identical under both settings so that the first shortage event may be either 2 or 4. The sixth column in Table 1 demonstrates that the relation between the usable inventory levels under both settings may be either $I_j^{\Omega}(I_0, B_0+1) = I_j^{\Omega}(I_0, B_0)$ or $I_j^{\Omega}(I_0, B_0+1) = I_j^{\Omega}(I_0, B_0) - 1$. In the latter case, the subsequent shortage event may be either 1 or 3. Resulting again in the same possible inventory relations.

It is noticeable from the fourth and fifth columns of Table 1 that whenever a shortage occurs under setting (I_0, B_0) it also occurs under setting (I_0, B_0+1), but not the opposite. That is, for any occurrence of shortage, we have $\Delta_j^{\Omega}(I_0, B_0+1) \geq \Delta_j^{\Omega}(I_0, B_0)$ and $\Theta_j^{\Omega}(I_0, B_0+1) \geq \Theta_j^{\Omega}(I_0, B_0)$. In addition, for any demand occurrence where no shortage occurs, all indicators equal zero, and the inequalities hold trivially. By summing these inequalities for all demand occurrences and multiplying by the related penalties we obtain for the set Ω: $F^{\Omega}(I_0, B_0+1) \geq F^{\Omega}(I_0, B_0)$

Theorem 1: The EUDF $F(I_0, B_0)$ is non-decreasing in the initial inventory of unusable bicycles B_0.

Proof: Consider the shortage occurrences given two initial settings of a station: (I_0, B_0+1) and (I_0, B_0). We will show that $F^{R}(I_0, B_0+1) \geq F^{R}(I_0, B_0)$, for any demand realization R and thus claim that it holds for the expectation. For a given demand realization R, we divide the set of demand occurrences E^{R} to sequences of demand occurrences such that in each sequence there are no return attempts of unusable bicycles and the sequences are separated by return attempts of unusable bicycles. Note that for the first sequence the conditions are as in Lemma 1, i.e. the inequality holds. Following this sequence there are three different possibilities: (i) it is possible to return the unusable bicycle under both settings. (ii) it is not possible to return the unusable bicycle under both settings. (iii) it is possible to return the unusable bicycle under setting (I_0, B_0) but not under (I_0, B_0+1). When (i) or (ii) occurs, the difference in the total number of shortages up to this point, between the two settings remains unchanged and the same analysis may be repeated for the next sequence, since there is still a difference of one unusable bicycle between the two. After (iii) occurs the number of usable and unusable bicycles are identical under both settings, therefore, from this point and on the station faces exactly the same shortages under both initial settings. Now, since this is true for any demand realization it is also true for the expectation, thus we obtain: $F\left(I_0, B_0+1\right) \geq F\left(I_0, B_0\right)$.

Remark: Since the EUDF $F\left(I_0, B_0\right)$ is non-decreasing in B_0, the function is minimized at $B_0=0$, as expected. However, due to time and capacity constraints, the operator may not be able to remove all unusable bicycles (or even visit all stations in which there are unusable bicycles) therefore it is important to analyze the EUDF for all possible values of $\left(I_0, B_0\right)$.

Next, we prove the following three inequalities, which are needed for the convexity proof of the EUDF that will be presented in Section 3.2.

- $F\left(I_0, B_0+2\right)-F\left(I_0, B_0+1\right)-F\left(I_0, B_0+1\right)+F\left(I_0, B_0\right) \geq 0$
- $F\left(I_0+2, B_0\right)-F\left(I_0+1, B_0\right)-F\left(I_0+1, B_0\right)+F\left(I_0, B_0\right) \geq 0$
- $F\left(I_0+1, B_0+1\right)-F\left(I_0+1, B_0\right)-F\left(I_0, B_0+1\right)+F\left(I_0, B_0\right) \geq 0$

Observe that the first two inequalities mean that the EUDF is convex in each of the variables (I_0, B_0) independently. The proofs for these inequalities are given under the following assumption:

Assumption 1: No unusable bicycles are returned to the station during the given period.

While this assumption may seem restrictive, note that the probability that a returned bicycle is unusable is low (see, for example, the discussion about the maintenance reports of NYC

Bikeshare in Section 5). Hence, a major share of the effect of unusable bicycles is already captured by the unusable bicycles that are already parked in the station, i.e., in the initial state of the station. We note that without Assumption 1 it is possible to "cook" an example in which the EUDF is non-convex, see Appendix A3. However, note that the approximation method of the EUDF presented in Section 4 does not rely on Assumption 1. Moreover, in section 5, we evaluate the EUDF using real life demand data, including returns of unusable bicycles and confirm that the convexity conditions hold or, at worst, are violated with a negligible margin.

Lemma 2: Under Assumption 1, the EUDF $F\left(I_0, B_0\right)$ is convex in the initial inventory of unusable bicycles B_0, i.e.: $F\left(I_0, B_0+2\right)-F\left(I_0, B_0+1\right) \geq F\left(I_0, B_0+1\right)-F\left(I_0, B_0\right)$.

The proof of this Lemma is based on an approach similar to the one used in the proof of Lemma 1. For brevity of the main text, we present the complete proof in Appendix A4.

Lemma 3: under Assumption 1, the EUDF $F\left(I_0, B_0\right)$ is convex in the initial inventory of usable bicycles I_0, i.e.:. $F\left(I_0+2, B_0\right)-F\left(I_0+1, B_0\right) \geq F\left(I_0+1, B_0\right)-F\left(I_0, B_0\right)$.

We remark that the convexity proof provided in Raviv and Kolka (2013) can be used here since the effective capacity under all three settings is equal and remains constant during the entire given period. Here we prove this result through an alternative approach that is later used in the proof of Lemma 4.

Proof: Note that for each side of the inequality the station's initial setting varies only by the initial inventory of usable bicycles. Under Assumption 1, the number of unusable bicycles remains the same in all these settings. Therefore, in a pair of settings, once a shortage occurs in one of the settings, either for a bicycle or for a locker, the number of usable bicycles equalizes and from that point on the number of shortages are equal under both settings. For example, for settings (I_0+1, B_0) and (I_0, B_0) if the first shortage is for a bicycle, then the demand can be satisfied by setting (I_0+1, B_0) but not by setting (I_0, B_0) so that right after the shortage occurs, under both settings the station is empty. Similarly, if the first shortage is for a locker, then the demand can be satisfied under setting (I_0, B_0) but not under setting (I_0+1, B_0) so that right after the shortage occurs, under both settings the station is full. That is, for a given realization, the shortage difference between the two settings may be either -1,0 or 1. In Table 2 we compare the two sides of the inequality by exhibiting all possible combinations of first shortage occurrences for a given demand realization R. As can be seen, for all possible combinations we obtain $F^R(I_0+2, B_0) - F^R(I_0+1, B_0) \geq F^R(I_0+1, B_0) - F^R(I_0, B_0)$. Consequently, by summing over all demand realizations we obtain: $F(I_0+2, B_0) - F(I_0+1, B_0) \geq F(I_0+1, B_0) - F(I_0, B_0)$

Lemma 4: under Assumption 1, for the EUDF $F(I_0, B_0)$ the following inequality is maintained:

$$F(I_0+1, B_0+1) - F(I_0, B_0+1) \geq F(I_0+1, B_0) - F(I_0, B_0)$$

Proof: Observe again that in each side of the inequality the settings differ by one usable bicycle, therefore we can again compare the first shortage events (as in the proof of Lemma 3). In Table 3 we present the possible combinations of first shortage occurrences for the two sides of the inequality. Note that the last combination presented in Table 2 is not possible in this case and therefore does not appear in Table 3. It is observable from Table 3 that for all possible combinations we obtain $F^R(I_0+1, B_0+1) - F^R(I_0, B_0+1) \geq F^R(I_0+1, B_0) - F^R(I_0, B_0)$.

Consequently, by summing over all demand realizations we obtain: $F\left(I_0+1, B_0+1\right)-F\left(I_0, B_0+1\right) \geq F\left(I_0+1, B_0\right)-F\left(I_0, B_0\right)$.

3.2. Convexity analysis of the EUDF

Recall that the EUDF is a bivariate discrete function. While the concept and definition of discrete convexity of univariate functions is quite similar to continuous convexity, this is not the case for multivariate discrete functions. In fact, several different definitions of convexity are given in the literature for multivariate discrete functions. In Murota and Shioura (2001), Murota (2009) and Moriguchi and Murota (2011), several classes of multivariate discrete convex functions are defined and the relationship among these classes is presented. We next outline some of these definitions and then prove that under Assumption 1, the EUDF is contained in these classes.

Definition 1: Convex extensibility (Murota 2009)

A function $f:\mathbb{Z}^n \rightarrow \mathbb{R}$ is said to be *convex-extensible* if there exists a convex function $\bar{f}:\mathbb{R}^n \rightarrow \mathbb{R}$ such that $\bar{f}(x)=f(x)$ for all $\mathrm{x} \in \mathbb{Z}^{\mathrm{n}}$.

Definition 2: $M^{\natural}$-convex (based on Moriguchi and Murota 2011)

Denote the i^{th} unit vector by e_i and $e_0=0$, denote the domain of f by $\operatorname{dom} f=\left\{x \in \mathbb{Z}^n \mid f(x)<+\infty\right\}$ and denote the positive and negative supports of a vector x by:

$$\operatorname{supp}^{+}(x)=\left\{i \in\{1, \ldots, n\} \mid x_i \rangle 0\right\}$$

$$\operatorname{supp}^{-}(x)=\left\{i \in\{1, \ldots, n\} \mid x_i<0\right\}$$

A function $f:\mathbb{Z}^n \rightarrow \mathbb{R}$ is $M^{\natural}$-*convex* if it satisfies the following exchange property:

(M$^\natural$-EXC) $\forall x, y \in \operatorname{dom} f$, $\forall i \in \operatorname{supp}^{+}(x-y), \exists j \in \left(\operatorname{supp}^{-}(x-y) \cup \{0\}\right)$ such that

$$f(x)+f(y) \geq f\left(x-e_i+e_j\right)+f\left(y+e_i-e_j\right).$$

See Murota and Shioura (2001) for a further discussion on M$^\natural$-convexity.

Definition 3: Discrete Hessian matrix (Moriguchi and Murota 2011)

The discrete Hessian $H(x)=\left(H_{ij}(x)\right)$ of $f:\mathbb{Z}^n \to \mathbb{R}$ at $x \in \mathbb{Z}^n$ is defined by

$$H_{ij}(x)=f\left(x+e_i+e_j\right)-f\left(x+e_i\right)-f\left(x+e_j\right)+f(x)$$

Definition 4: (Theorem 3.1 in Moriguchi and Murota 2011)

A function $f:\mathbb{Z}^n \to \mathbb{R}$ is $M^\natural$*-convex* if and only if the discrete Hessian matrix $H(x)$ in Definition 3 satisfies the following conditions for each $x \in \mathbb{Z}^n$:

(i) $H_{ij}(x) \geq \min\left(H_{ik}(x), H_{jk}(x)\right)$ *if* $\{i,j\} \cap \{k\} = \varnothing$

(ii) $H_{ij}(x) \geq 0$ *for any* (i,j).

Note that $H_{ii}(x) \geq 0$ means that f is convex in the variable i.

Theorem 2: (Follows from Theorem 3.9 and Theorem 3.3 in Murota and Shioura 2001)

An $M^\natural$-convex function is convex-extensible.

Next, given the above definitions, we present and prove the main theorem of this study:

Theorem 3: Under Assumption 1, the EUDF $F(I_0, B_0)$ is $M^\natural$-convex.

Proof: Since the EUDF is a bivariate function, condition (i) of Definition 4 is not relevant for our analysis, and condition (ii) of Definition 4 reduces to the three inequalities that were presented and proved in Section 3.1. Given the proofs of Lemmas 2-4, the EUDF satisfies the conditions given in Definition 4. Therefore the discrete Hessian of the EUDF is positive semidefinite and the EUDF is $M^{\natural}$-convex

Note that since the EUDF is $M^{\natural}$-convex we conclude by Theorem 2 that it is also convex-extensible. Therefore, there exists a continuous convex function that has identical values at all integer points in the range of the EUDF. In the next section we present a method to approximate the EUDF by a convex polyhedral function that has the same values at integer points. Under Assumption 1, the EUDF is convex-extensible and therefore the approximation will provide an exact description of the function. More importantly, the results of the numerical experiment that will be presented in Section 5, demonstrate that even if Assumption 1 is relaxed, the approximation is very accurate.

4. A convex polyhedral function approximation

The approximation procedure of the EUDF is divided into two steps. In the first step we approximate the values of the EUDF for each possible combination of integer initial inventory levels $\left(I_0, B_0\right)$. In the second step an LP model is used to fit a convex polyhedral function to the values calculated in the first stage. That is, the epigraph of the EUDF is defined, approximately, as an intersection of half spaces.

Recall that the EUDF is the expectation of all possible demand realizations. One approach for estimating the expectations is by using Monte Carlo simulation. However, this process may require long calculation times and can be very noisy. Moreover since this calculation needs to be carried out for each possible initial setting and for every station in the bike sharing system, this approach seems impractical. Instead, we adopt an approximation approach which is similar to the one presented in Raviv and Kolka (2013). This approach is based on a representation of the states of the station along the given period as a continuous time Markov chain.

To that end, we assume that the arrival processes of renters and returners to the station are time heterogeneous Poisson processes, with arrival rates μ_t and λ_t, respectively. When a user returns a bicycle to the station there is a probability ϕ that the bicycle is unusable. That is, the returning rate of usable bicycles at time period t is $(1-\phi)\lambda_t$ and the returning rate of unusable bicycles at time period t is $\phi\lambda_t$. Since the arrival processes of renters and returners reflect the arrivals of many independent users, we believe that this Markovian model is an adequate description of reality.

Recall that during the given period, no repositioning activities are being executed. While the inventory level of usable bicycles may increase or decrease along the day, the inventory level of unusable bicycles may only increase. Therefore, for any $\phi > 0$, in steady-state, the station will be full with unusable bicycles. However, we are interested in analyzing the dynamics of the station rather than its steady-state. Moreover, a station that is regulated in a sufficient manner is not likely to reach its steady-state. A description of the continuous-time Markov chain that represents the dynamic of the station is given in Figure 1.

Let $\pi_{(I_0,B_0),(I,B)}(t)$ denote the probability that the station is in state (I,B) at time t given that in time 0 it was in state (I_0,B_0). Now, it is possible to state the EUDF in terms of the transition probabilities as follows:

$$F(I_0,B_0)=\int_0^T\left(\left(\sum_{k=B_0}^{C}\pi_{(I_0,B_0),(0,k)}(t)\right)\mu_t p+\left(\sum_{k=B_0}^{C}\pi_{(I_0,B_0),(C-k,k)}(t)\right)\lambda_t h\right)dt \quad (2)$$

The first term in the integral represents the user dissatisfaction due to bicycle shortages. It is calculated by the probability that at time t the station is empty, multiplied by the renting rate μ_t and the penalty for bicycle shortage p. The second term represents the user dissatisfaction due to locker shortages. It is calculated by the probability that at time t the station is full, multiplied by the returning rate λ_t and the penalty for locker shortage h. The evaluation of (2) is numerically obtained by discretizing the integral to short intervals of length d and calculating the following sum:

$$F\left(I_0,B_0\right)=d\sum_{i=1}^{T/d}\left(\left(\sum_{k=B_0}^{C}\pi_{(I_0,B_0),(0,k)}\left(\left(i-0.5\right)d\right)\right)\mu_t p+\left(\sum_{k=B_0}^{C}\pi_{(I_0,B_0),(C-k,k)}\left(\left(i-0.5\right)d\right)\right)\lambda_t h\right)$$

It is assumed that T/d is an integer. The value of $\pi_{(I_0,B_0),(I,B)}\left(t\right)$ is numerically evaluated for each of the T/d points in time by the method presented in Raviv and Kolka (2013), we refer the reader to Section 4 in their paper.

Next we discuss the fitting of a convex polyhedral function to the approximated values of the EUDF. As the EUDF is convex-extensible (under Assumption 1), there exists a continuous convex function that has identical values as the EUDF in all integer points. We denote this function by $f\left(I_0,B_0\right)$. Though this function is unknown, we can use the fact that it is convex. First, let us state the following proposition:

Proposition 1: Supporting a convex function (adapted from Proposition 2.6.2 in Ben-Tal and Nemirovski 2013)

For any point $\bar{\mathrm{x}}$ in the domain of a convex function $f:\mathbb{R}^n\rightarrow\mathbb{R}$ there exists an affine function $f_{\bar{\mathrm{x}}}\left(\mathrm{x}\right)=\mathrm{a}^{\mathrm{T}}\mathrm{x}+\mathrm{b}$, such that $f_{\bar{\mathrm{x}}}\left(\bar{\mathrm{x}}\right)=f\left(\bar{\mathrm{x}}\right)$ and $f_{\bar{\mathrm{x}}}\left(\mathrm{x}\right)\le f\left(\mathrm{x}\right)$ for all $\mathrm{x}\in\mathbb{R}^n$.

Let θ be the range of the EUDF, namely $\theta = \left\{ \left(I_0, B_0 \right) \in \mathbb{Z}^2 | I_0 \geq 0, B_0 \geq 0, I_0 + B_0 \leq C \right\}$. Since $f\left(I_0, B_0\right)$ is convex and given Proposition 1, for each point $\left(\tilde{I}_0, \tilde{B}_0\right) \in \theta$ there exists a plane that satisfies:

$$\alpha_{\left(\tilde{I}_0, \tilde{B}_0\right)} I + \beta_{\left(\tilde{I}_0, \tilde{B}_0\right)} B + \gamma_{\left(\tilde{I}_0, \tilde{B}_0\right)} = f\left(\tilde{I}_0, \tilde{B}_0\right) \text{ (3)}$$

and

$$\alpha_{\left(\tilde{I}_0, \tilde{B}_0\right)} I_0 + \beta_{\left(\tilde{I}_0, \tilde{B}_0\right)} B_0 + \gamma_{\left(\tilde{I}_0, \tilde{B}_0\right)} \leq f\left(I_0, B_0\right) \ \forall \left(I_0, B_0\right) \in \theta \text{ (4)}$$

By generating a supporting plane for each point $(I_0, B_0) \in \theta$, we obtain the following convex polyhedral function:

$$\breve{f}\left(I_0, B_0\right) = \max_{\left(\tilde{I}_0, \tilde{B}_0\right) \in \theta} \left(\alpha_{\left(\tilde{I}_0, \tilde{B}_0\right)} \cdot I_0 + \beta_{\left(\tilde{I}_0, \tilde{B}_0\right)} \cdot B_0 + \gamma_{\left(\tilde{I}_0, \tilde{B}_0\right)} \right)$$

Note that for each point $(I_0, B_0) \in \theta$ we have $\breve{f}\left(I_0, B_0\right) = f\left(I_0, B_0\right) = F\left(I_0, B_0\right)$.

However, the calculation of a plane that satisfies (3) and (4) may, in some cases, be impossible due to the following reasons: (i) when Assumption 1 is relaxed, the EUDF is not necessarily convex-extensible. (ii) in the calculation of the approximated EUDF values, numerical errors may occur. To overcome these issues, we construct a plane for point $\left(\tilde{I}_0, \tilde{B}_0\right)$ such that (4) is satisfied but some error is allowed in (3). Namely, the constructed plane may pass under $F\left(\tilde{I}_0, \tilde{B}_0\right)$. Our goal is to construct a plane that passes as close as possible to $F\left(\tilde{I}_0, \tilde{B}_0\right)$. We use the following LP formulation to achieve this goal:

Decision variables:

α, β, γ Coefficients of the fitted plane

s Gap at the point $\left(\tilde{I}_0, \tilde{B}_0\right)$

Model:

$$Minimize \quad s \quad (5)$$

$$s.t. \alpha \cdot I_0 + \beta \cdot B_0 + \gamma \leq F\left(I_0, B_0\right) \qquad \forall \left(I_0, B_0\right) \in \theta \setminus \left\{\left(\tilde{I}_0, \tilde{B}_0\right)\right\} \quad (6)$$

$$\alpha \cdot \tilde{I}_0 + \beta \cdot \tilde{B}_0 + \gamma = F\left(\tilde{I}_0, \tilde{B}_0\right) - s \qquad (7)$$

$$\alpha, \beta, \gamma \quad free \quad (8)$$

$$s \geq 0 \quad (9)$$

The objective function (5) minimizes the gap between the fitted plane and the EUDF at point $\left(\tilde{I}_0, \tilde{B}_0\right)$. Constraint (6) requires that the fitted plane pass under the EUDF at all other integer points. Constraint (7) defines the gap between the plane and the EUDF at point $\left(\tilde{I}_0, \tilde{B}_0\right)$. In Constraints (8)-(9) the definitions of the decision variables are given.

It is possible to redefine $\breve{f}\left(I_0, B_0\right)$ with the values of $\alpha_{\left(\tilde{I}_0, \tilde{B}_0\right)}, \beta_{\left(\tilde{I}_0, \tilde{B}_0\right)}$ and $\gamma_{\left(\tilde{I}_0, \tilde{B}_0\right)}$ obtained by the LP above for all $\left(I_0, B_0\right) \in \theta$. The maximal value of s over all the points $\left(\tilde{I}_0, \tilde{B}_0\right)$ is an upper bound on the gap between $\breve{f}\left(I_0, B_0\right)$ and $F\left(I_0, B_0\right)$. In cases where the gap is zero for all constructed planes, we can say that the EUDF is convex-extensible.

5. Numerical results

In this section, we evaluate the accuracy of the polyhedral approximation of the EUDF and derive some insights regarding the effect of unusable bicycles on user dissatisfaction. We analyze two case studies, the Washington D.C. bike sharing system, Captial Bikeshare, and the New-York bike sharing system, Citi Bike. In Table 4, we present some statistics regarding the two systems. The data can be downloaded from the systems' websites: http://www.capitalbikeshare.com/trip-history-data and https://www.citibikenyc.com/system-data. Using this data we have estimated the renting and returning rates on weekdays in each station for each interval of 30 minutes for a period of 24 hours ($T=1$) starting at midnight. The processed renting rates, return rates and the station capacities can be downloaded from: http://www.eng.tau.ac.il/~morkaspi/publications.html.

The approximation of the EUDF was coded in MathWorks Matlab™. The plane fitting LP model was solved using IBM ILOG CPLEX Optimization Studio 12.6. The procedure was tested on an Intel Core i7 desktop. For Capital BikeShare (resp., Citi Bike), on average, the entire process of calculating the EUDF took three seconds (resp., 70 seconds) per station and the polyhedral function fitting took less than a second (5 seconds) per station. This means that approximating the function for all stations as an input for a repositioning and collection optimization can be executed in acceptable time. Moreover, since the calculation for each station is done independently, the calculation procedure is amendable for parallelization. In addition, the renting/returning rates are typically not estimated on a daily basis and therefore the polyhedral functions will not be updated very often.

For each station, we have approximated the EUDF by fine discretization to intervals of one minute. This was done for varying values of, ϕ, the probability of a bicycle to be returned unusable, in the range 0%-5% in increments of 1%. In addition, to represent extreme scenarios, we have also set the probability to the unrealistic value of 20%. Three pairs of shortage penalties were tested: $(p=1, h=1)$, $(p=0.75, h=1.25)$, $(p=0.5, h=1.5)$. In the first case, the penalty for both types of shortages is the same, in the second (resp., third) case the penalty for locker shortages is 1.66 (resp., 3) times the penalty for bicycle shortages. Note that we change both penalty weights simultaneously, in order to keep the sum of the penalty weights fixed. This allows, to some extent, a comparison between different penalty weight configurations. The EUDF evaluation for all stations of the studied systems, are available at http://www.eng.tau.ac.il/~morkaspi/publications.html. For each station and each polyhedral function, we calculated the maximal absolute gap with respect to the approximated values of the EUDF and the maximal relative gap over all possible points, where:

$$Ralative\ gap = \frac{absolute\ gap}{EUDF\ value}$$

In Table 5, we present the aggregated values for all 232 stations of Capital BikeShare and all 332 stations of Citi Bike. In the first column, we present the shortage penalties. The second column presents the probability for a bicycle to be returned unusable. The third to fifth columns present for the Capital Bikeshare system, the number of stations in which there was no gap at all in fitting the convex polyhedral function and the maximal absolute and relative gaps over all stations. The sixth to eighth columns present these figures for the Citi Bike system.

One can observe in Table 5 that for almost all the penalty settings (p, h) and the probabilities (ϕ) that we examined the EUDF could be described exactly by a convex polyhedral function. In the few cases where such a description is not exact, its maximal gap (error) is negligible. That is, the EUDF is approximated very accurately by a convex polyhedral function and this accuracy is not highly sensitive to the values of p, h and ϕ. Indeed, for any practical purpose, the convex polyhedral function can be considered as an exact description of the approximated EUDF. To test the sensitivity of the approximation to the period length, we repeated the above analysis for $T = 3$. That is, we estimated the expected shortages assuming that during three days the inventory levels are not externally altered. Similar results were obtained; for the sake of brevity we present these results in Appendix A5, see Table 10.

Recall that in the EUDF approximation presented in Section 4, we assumed that the arrival processes of renters and returners to the stations are time heterogeneous Poisson processes. To examine this assumption, we conducted Chi-Square goodness of fit tests to check the hypothesis that over each 30-minute interval, the number of arriving renters to each of the stations is Poisson distributed. The data was collected over 65 working days (n=65) in the Capital Bikeshare case, and 44 working days (n=44) in the Citi Bike case. Since we examined each interval with non-zero demand (aggregately over the entire period) we conducted more than 24,000 tests in the two systems. Our null hypothesis was rejected (with alpha=0.05) in about 13.5% of the cases. The conclusion from this inquiry is that a non-homogenous Poisson arrival process, with rates that vary every half an hour, is a good approximation for the demand process in most stations most of the time. Nevertheless, the actual arrival process may sometimes deviate significantly from this convenient working assumption, which could possibly result in some cases from the noise in the data.

To investigate this point more thoroughly, we examined the sensitivity of our results to the Poisson assumption. In particular, we studied the effect of bulk arrivals of renters and returners on the accuracy of the convex polyhedral approximation. A bulk arrival process was artificially created in order to demonstrate a reasonable (yet hypothetical) demand process. This process has the same average rate of arrivals during each half an hour interval at each station as the original Poisson arrival process. However, the two processes are different in nature. In Appendix A6, we explain how the approximation method is adapted to represent arrival of bulks and present the numerical results of using the convex polyhedral function to describe it for our two case studies. We conclude that taking into account explicitly the arrival of bulks results with an even more accurate polyhedral approximation.

During the above examination, we observed that the number of bicycles that minimizes the EUDF under Poisson arrivals is often very similar to the one that minimizes it under the bulk arrival process. To verify this observation, we compared the optimal initial inventories under both processes in all 564 stations of Capital Bikeshare and CitiBike with seven different levels of ϕ and three combinations of penalty weights (as listed in Table 5 and Table 12). Overall, 11,844 comparisons were made. In about 98.6% of the cases, the prescribed optimal initial inventory was either identical or with a difference of up to two bicycles. We interpret the results of this inquiry as another practical justification for using the working assumption of Poisson arrival processes. Indeed, this is a common practice in the analysis of many service systems, mainly because it is easy to estimate the parameters of this stochastic process from demand data.

The observation that the quality of the convex polyhedral approximation is satisfactory under a diverse set of arrival processes strengthens our belief that in real life scenarios the EUDF is convex-extensible even when Assumption 1 is relaxed. Moreover, recall that in the case $\phi = 0$, the EUDF is $M^{\natural}$-convex (Theorem 3), and therefore a polyhedral convex function can be fitted with no gap. Nevertheless, even for $\phi = 0$, small gaps are observed in Table 5, Table 10 and Table 12 (in Appendix A5) as a result of numerical errors. This fact indicates that also for $\phi > 0$ significant portions of the gaps may originate from numerical errors in the approximation of the EUDF values and not from the true structure of the function. For the very general time heterogeneous Poisson arrival and bulk arrival processes of renters and returners, it remains an open question whether the EUDF is indeed convex-extensible and if not, whether the approximation errors can be bounded.

In Table 5, the range of probabilities we examined was 0%-20%. In order to estimate the probability that a bicycle is returned unusable in a real system, relevant information should be collected. In Citi Bike, the total number of trips taken in 2014 was 8,791,987 and the total number of bicycle repairs was 34,806. Therefore, a reasonable estimator of ϕ is about 0.004 (0.4%). The results provided in Table 5 demonstrate that for such a probability, the approximation of a polyhedral convex function is very accurate.

Finally, we examine the effect of two aspects of the presence of unusable bicycles on the user dissatisfaction. First, unusable bicycles that are parked in a station, block lockers and virtually reduce the effective capacity of the station. Second, inaccurate estimation of the number of unusable bicycles may lead to inferior operational decisions. Focusing on a single station, we present in Table 6, as an example, the approximated EUDF values for a station with 32 lockers. Each value in the table is the EUDF calculated for a given combination of initial number of usable and unusable bicycles. The number of usable bicycles is presented on the rows of the table and the number of unusable on the columns. The values are calculated for $\phi = 1\%$ and $h = p = 1$. We note that for this example, the EUDF could be approximated by a polyhedral function without any errors, that is, the function is convex-extensible.

Note that moving from left to right on the columns of Table 6 is equivalent to decreasing the effective capacity of the station. It is observable that such a decrease results with an increase of the user dissatisfaction. In addition, the differences between each pair of consecutive columns increase as the effective capacity decreases. This demonstrates what we have previously proved in Section 3, that the EUDF is convex and monotonically increasing in the initial number of unusable bicycles (Theorem 1, Lemma 2).

Inaccurate estimation of the number of unusable bicycles or disregarding of the presence of unusable bicycles may lead to discrepancies in user dissatisfaction estimation. For example, if the initial inventory of bicycles is set to 16 (half of the station capacity) and are all usable, the expected daily number of shortages is 7.96. However, if three of these bicycles were actually unusable, the expected number of shortages would be 9.89, that is an increase of 24% in the user dissatisfaction. Note that all values in this table are evaluated with the assumption that until the next review of the station, additional unusable bicycles may be returned to the station ($\phi = 1\%$). However, disregarding this possibility, may lead to additional discrepancies. In this example, if the initial inventory is set to 16 assuming all bicycles are usable and that all the bicycles that would be returned to the station would be usable ($\phi = 0$), the expected number of shortages would be 7.07, that is, a further underestimation.

The above effects emphasize the need for having correct information regarding unusable bicycles in order to plan the replenishment and repositioning activities properly. Furthermore, collecting unusable bicycles may have a greater effect on the service level at a station as compared to adding/removing usable bicycles and therefore should be prioritized.

6. Conclusions

In this study, we have extended the user dissatisfaction function to account for the number of unusable bicycles in addition to the number of usable bicycles. The presence of unusable bicycles significantly affects the quality of service given to the users of bike sharing systems. Maintenance aspects in bike sharing are studied for the first time in this paper and in Kaspi et al. (2016a, 2016b).

We have demonstrated that a convex polyhedral function can be accurately fitted to the EUDF. In particular, we have proved that the EUDF is $M^{\natural}$-convex (and thus convex-extensible) for $\phi = 0$. The results of our numerical experiment suggest that in real life setting the EUDF is convex-extensible also for $\phi > 0$. As a consequence, the EUDF can be used in linear optimization models for planning of the operational activities. For example, we present in Kaspi (2015, pp-105-108) an integrated optimization model for bicycle repositioning and collection of unusable bicycles. In that model the EUDF appears in the objective function, representing the quality of service given to the users. Its property of convex extensibility enables the exact representation of all of its relevant integer points, by adding to the formulation linear constraints that support this function. In addition, the extended user dissatisfaction model can assist in strategic planning, e.g., deciding on the size of the bicycle fleet, capacity of the stations, manpower requirement for operations and maintenance activities, etc.

The numerical results demonstrate that the presence of unusable bicycles may highly increase user dissatisfaction. Thus, even though only a small fraction of the bicycles is returned unusable the effect of these bicycles is significant. Therefore, this matter should receive more attention in the planning process. Particularly, system operators should invest resources in detection and collection of unusable bicycles. Accurate information regarding bicycle usability should be obtained and made available to the operators and the users of the system. The former can use it to optimize the maintenance and repositioning activities and the latter to better plan their itineraries.

Acknowledgements

The valuable comments of three anonymous reviewers, the Area Editor and the Department Editor, Professor Enver Yucesan, are gratefully acknowledged by the authors.

References:

Angeloudis, P., Hu, J., Bell, M-G-H, 2012. A strategic repositioning algorithm for bicycle-sharing schemes. In: *Transportation Research Board 91st Annual Meeting Compendium of Papers*.

Benchimol, M., Benchimol, P., Chappert, B., Taille, A-D-L., Laroche, F., Meunier, F., Robinet, L,. 2011. Balancing the stations of a self-service "Bike Hire" system. *RAIRO Operations Research*, 45, 37–61.

Ben-Tal, A., Nemirovski A., 2013. Optimization III. Georgia Institute of Technology. http://www2.isye.gatech.edu/~nemirovs/OPTIII_LectureNotes.pdf

Castro, F.A. 2011. The contribution of bike-sharing to sustainable mobility in Europe, Doctoral Thesis, Technische Universität Wien.

Chemla, D., Meunier, F., Wolfer-Calvo, R., 2013a. Bike sharing systems: solving the static rebalancing problem. *Discrete Optimization*,10(2), 120-146.

Chemla, D., Meunier, F., Pradeau, T., Wolfler Calvo, R. and Yahiaoui, H. 2013b. Self-Service Bike Sharing Systems: Simulation, Repositioning, Pricing. Working paper.

http://hal.archives-ouvertes.fr/docs/00/82/40/78/PDF/RealTime-BikeSharing_final.pdf.

Contardo, C., Morency, C., Rousseau, L-M., 2012. Balancing a dynamic public bike sharing system. https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2012-09.pdf

Dell'Amico, M., Hadjicostantinou, E., Iori, M., & Novellani, S. (2014). The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega*, *45*, 7-19

DeMaio, P., Meddin, R., (2009). The Bike-sharing blog. http://bike-sharing.blogspot.co.il/2009/07/kickin-it-into-high-gear-this-summer-in.html

DeMaio, P., Meddin, R., (2014). The Bike-sharing blog. http://bike-sharing.blogspot.co.il/2014/06/the-bike-sharing-world-first-week-of.html

DeMaio, P., Meddin, R., (2015). The Bike Sharing World Map. http://www.bikesharingmap.com/

Erdoğan, G., Battarra, M., & Calvo, R. W. (2015). An exact algorithm for the static rebalancing problem arising in bicycle sharing systems. *European Journal of Operational Research*, *245*(3), 667-679.

Erdoğan, G., Laporte, G., & Calvo, R. W. (2014). The static bicycle relocation problem with demand intervals. *European Journal of Operational Research*, *238*(2), 451-457.

Forma, I. A., Raviv, T., & Tzur, M. 2015. A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation Research Part B: Methodological*, *71*, 230-247.

Fricker, C., & Gast, N. 2014. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *EURO Journal on Transportation and Logistics*, 1-31.

George, D-K., Xia, C-H., 2011. Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. *European Journal of Operational Research*, 211(1), 198-217.

Kaspi, M. 2015. Bike-sharing systems: parking reservation policies and maintenance operations (doctoral dissertation). Tel-Aviv University.

http://www.eng.tau.ac.il/~morkaspi/publications.html

Kaspi, M., Raviv, T., & Tzur, M. 2014. Parking reservation policies in one-way vehicle sharing systems. *Transportation Research Part B: Methodological*, *62*, 35-50.

Kaspi, M., Raviv, T., & Tzur, M. 2016a, Collection of Unusable Bicycles in Bike-Sharing Systems. Working paper.

Kaspi, M., Raviv, T., & Tzur, M. 2016b, Detection of Unusable Bicycles in Bike-Sharing Systems. Omega, forthcoming.

Kaspi, M., Raviv, T., Tzur, M., & Galili, H. 2016c. Regulating vehicle sharing systems through parking reservation policies: Analysis and performance bounds. European Journal of Operational Research, 251(3), 969-987.

Kloimüllner, C., Papazek, P., Hu, B. and Raidl, G.R. (2014). Balancing Bicycle Sharing Systems: An Approach for the Dynamic Case. In: C. Blum and G. Ochoa (Eds.): EvoCOP 2014, LNCS 8600, 73–84.

Lin, J-R., Yang, T-H., 2011. Strategic design of public bicycle sharing systems with service level constraints. *Transportation Research Part E*, 47, 284–294.

Lin, J. R., Yang, T. H., & Chang, Y. C. (2013). A hub location inventory model for bicycle sharing system design: Formulation and solution. *Computers & Industrial Engineering*, 65(1), 77-86.

Moriguchi, S., Murota, K. (2011). Note on Discrete Hessian Matrix and Convex Extensibility. METR 2011-27, Department of Mathematical Informatics, University of Tokyo.

Murota, K. (2009). Recent developments in discrete convex analysis. In Research Trends in Combinatorial Optimization (pp. 219-260). Springer Berlin Heidelberg.

Murota, K., Shioura, A. (2001). Relationship of M-/L-convex functions with discrete convex functions by Miller and Favati–Tardella. *Discrete Applied Mathematics*, 115(1), 151-176.

Nair, R., Miller-Hooks, E., 2011. Fleet management for vehicle sharing operations. *Transportation Science*, 45(4), 524–540.

Pessach, D., Raviv, T., Tzur, M., 2014. Dynamic Repositioning in a Bike Sharing System. Working Paper, Tel-Aviv University.

Pfrommer, J., Warrington, J., Schildbach, G., & Morari, M. (2014). Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4), 1567-1578.

Raviv, T., Kolka, O., 2013. Optimal Inventory Management of a Bike Sharing Station. *IIE Transactions*, 45(10), 1077-1093.

Raviv, T., Tzur, M., Forma, I., 2013. The static repositioning problem in a bike sharing system: Models and Solution Approaches. *EURO Journal of Transportation and Logistics*, 2(3), 187-229.

Schuijbroek, J., Hampshire, R., van Hoeve, W-J., 2013. Inventory Rebalancing and Vehicle Routing in Bike Sharing Systems. *Tepper School of Business.* Paper 1491. http://repository.cmu.edu/tepper/1491

Shu, J., Chou, M., Liu, Q., Teo, C-P., Wang, I-L., 2013. Models for Effective Deployment and Redistribution of Bicycles within Public Bicycle-Sharing Systems. *Operations Research*, 61(6), 1346-1359.

Vogel, P., Saavedra, B. A. N., & Mattfeld, D. C. (2014). A hybrid metaheuristic to solve the resource allocation problem in bike sharing systems. In *Hybrid Metaheuristics* (pp. 16-29). Springer International Publishing.

Table 1: Shortage events in two settings with initial inventories $\left(\boldsymbol{I}_0, \boldsymbol{B}_0+1\right)$ and $\left(\boldsymbol{I}_0, \boldsymbol{B}_0\right)$

Shortage event	Usable bicycles before the shortage occurs	Shortage type	$\left(I_0, B_0+1\right)$	$\left(I_0, B_0\right)$	Usable bicycles after the shortage occurs
1	$I_{j-1}^{\Omega}\left(I_0, B_0+1\right)=I_{j-1}^{\Omega}\left(I_0, B_0\right)$	Bicycle	$\Delta_j^{\Omega}\left(I_0, B_0+1\right)=1$	$\Delta_j^{\Omega}\left(I_0, B_0\right)=$	$I_j^{\Omega}\left(I_0, B_0+1\right)=I_j^{\Omega}\left(I_0, B_0\right)$
2	$I_{j-1}^{\Omega}\left(I_0, B_0+1\right)=I_{j-1}^{\Omega}\left(I_0, B_0\right)$	Bicycle	$\Delta_j^{\Omega}\left(I_0, B_0+1\right)=1$	$\Delta_j^{\Omega}\left(I_0, B_0\right)=1$	$I_j^{\Omega}\left(I_0, B_0+1\right)=I_j^{\Omega}\left(I_0, B_0\right)$
3	$I_{j-1}^{\Omega}\left(I_0, B_0+1\right)=I_{j-1}^{\Omega}\left(I_0, B_0\right)$	Locker	$\Theta_j^{\Omega}\left(I_0, B_0+1\right)=1$	$\Theta_j^{\Omega}\left(I_0, B_0\right)=$	$I_j^{\Omega}\left(I_0, B_0+1\right)=I_j^{\Omega}\left(I_0, B_0\right)$
4	$I_{j-1}^{\Omega}\left(I_0, B_0+1\right)=I_{j-1}^{\Omega}\left(I_0, B_0\right)$	Locker	$\Theta_j^{\Omega}\left(I_0, B_0+1\right)=1$	$\Theta_j^{\Omega}\left(I_0, B_0\right)=$	$I_j^{\Omega}\left(I_0, B_0+1\right)=I_j^{\Omega}\left(I_0, B_0\right)$

Table 2: Possible combinations of first shortage occurrences

$F^R(I_0+2,B_0)-F^R(I_0+1,B_0)$		$F^R(I_0+1,B_0)-F^R(I_0,B_0)$	
First shortage occurrence	Difference	First shortage occurrence	Difference
Bicycle	$-p$	Bicycle	$-p$
Locker	h	Bicycle	$-p$
Locker	h	Locker	h
Locker	h	None	0
None	0	None	0
None	0	Bicycle	$-p$

Table 3: Possible combinations of first shortage occurrences

$F^R(I_0+1,B_0+1)-F^R(I_0,B_0+1)$		$F^R(I_0+1,B_0)-F^R(I_0,B_0)$	
First shortage occurrence	Difference	First shortage occurrence	Difference
Bicycle	$-p$	Bicycle	$-p$
Locker	h	Bicycle	$-p$
Locker	h	Locker	h
Locker	h	None	0
None	0	None	0

Table 4: Case studies statistics

	Capital BikeShare	**Citi Bike**
Location	Washington DC	New-York city
Transactions period	April-June 2013	July-August 2014
Number of stations	232	332
Average station capacity	16.6	34.5
Station capacities range	10-40	11-67
Average daily number of trips	7,822	33,296
Average number of rents/returns per station per day	33.7	100.3
Maximal number of rents (returns) per day in a station	203 (215)	501 (354)

Table 5: Numerical results for Capital Bikeshare and Citi Bike systems $(T = 1)$

Penalty Weights	Probability that a bicycle is returned unusable	Capital Bikeshare: Number of stations with no gaps (out of 232)	Capital Bikeshare: Expected daily bicycle and locker shortages ($T = 1 day$): Maximal absolute gap	Capital Bikeshare: Expected daily bicycle and locker shortages ($T = 1 day$): Maximal relative gap	Citi Bike: Number of stations with no gaps (out of 332)	Citi Bike: Expected daily bicycle and locker shortages ($T = 1 day$): Maximal absolute gap	Citi Bike: Expected daily bicycle and locker shortages ($T = 1 day$): Maximal relative gap
$p = 1$, $h = 1$	0	210	0.000151	0.000014	263	0.001580	0.000111
	0.01	211	0.000153	0.000014	266	0.001531	0.000106
	0.02	209	0.000155	0.000014	266	0.001526	0.000095
	0.03	214	0.000157	0.000014	273	0.001521	0.000078
	0.04	214	0.000160	0.000014	273	0.001519	0.000073
	0.05	217	0.000161	0.000015	275	0.001515	0.000039
	0.20	206	0.000904	0.000101	158	0.006101	0.000304
$p = 0.75$, $h = 1.25$	0	210	0.000106	0.000013	264	0.001131	0.000148
	0.01	211	0.000107	0.000014	266	0.001128	0.000110
	0.02	211	0.000109	0.000014	265	0.001125	0.000102
	0.03	213	0.000112	0.000014	269	0.001121	0.000091
	0.04	214	0.000115	0.000014	267	0.001118	0.000075
	0.05	217	0.000120	0.000014	274	0.001115	0.000059
	0.20	205	0.001129	0.000121	157	0.006276	0.000297
$p = 0.5$, $h = 1.5$	0	209	0.000101	0.000014	264	0.000842	0.000115
	0.01	209	0.000097	0.000000	266	0.000805	0.000113
	0.02	208	0.000095	0.000000	267	0.000723	0.000109
	0.03	210	0.000094	0.000000	268	0.000721	0.000101
	0.04	208	0.000092	0.000000	266	0.000718	0.000091
	0.05	214	0.000091	0.000000	272	0.000715	0.000078
	0.20	207	0.001357	0.000154	163	0.006451	0.000295

Table 6: User dissatisfaction as a function of the initial usable and unusable bicycles in a station with 32 lockers

Initial number of usable bicycles	32	11.45										
	31	10.52	11.74									
	30	9.67	10.81	12.05								
	29	8.93	9.96	11.12	12.40							
	28	8.29	9.23	10.29	11.48	12.78						
	27	7.76	8.60	9.56	10.64	11.86	13.21					
	26	7.35	8.09	8.94	9.92	11.03	12.28	13.67				
	25	7.03	7.69	8.44	9.32	10.32	11.46	12.75	14.17			
	24	6.82	7.39	8.06	8.83	9.73	10.76	11.93	13.25	14.72		
	23	6.69	7.20	7.78	8.46	9.26	10.17	11.23	12.44	13.80	15.32	
	22	**6.66**	7.10	7.61	8.21	8.91	9.72	10.66	11.75	13.00	14.41	15.97
	21	6.70	**7.08**	**7.53**	8.06	8.67	9.39	10.22	11.19	12.32	13.61	15.07
	20	6.82	7.15	7.55	**8.00**	8.54	9.17	9.91	10.77	11.77	12.94	14.27
	19	7.01	7.30	7.64	8.04	**8.51**	9.07	9.71	10.47	11.36	12.40	13.61
	18	7.27	7.52	7.82	8.17	8.58	**9.06**	**9.63**	10.30	11.08	12.01	13.09
	17	7.59	7.81	8.07	8.38	8.73	9.16	9.65	**10.24**	10.93	11.75	12.71
	16	7.96	8.16	8.39	8.66	8.97	9.34	9.77	10.29	**10.90**	11.62	12.47
	15	8.40	8.57	8.77	9.00	9.28	9.60	9.98	10.44	10.97	**11.60**	**12.36**
	14	8.88	9.03	9.21	9.42	9.66	9.94	10.28	10.68	11.15	11.71	12.37
	13	9.42	9.55	9.70	9.89	10.10	10.35	10.65	11.00	11.41	11.91	12.50
	12	10.00	10.11	10.25	10.41	10.60	10.82	11.09	11.40	11.77	12.20	12.73
	11	10.62	10.72	10.85	10.99	11.16	11.36	11.59	11.87	12.20	12.59	13.05
	10	11.28	11.38	11.49	11.62	11.77	11.94	12.15	12.40	12.70	13.04	13.46
	9	11.99	12.07	12.17	12.29	12.42	12.58	12.77	12.99	13.26	13.57	13.95
	8	12.73	12.80	12.89	13.00	13.12	13.27	13.44	13.64	13.88	14.17	14.51
	7	13.50	13.57	13.65	13.75	13.86	13.99	14.15	14.34	14.56	14.82	15.13
	6	14.31	14.37	14.44	14.53	14.64	14.76	14.91	15.08	15.28	15.52	15.81
	5	15.14	15.20	15.27	15.35	15.45	15.56	15.70	15.86	16.05	16.28	16.55
	4	16.00	16.06	16.12	16.20	16.29	16.40	16.53	16.68	16.86	17.07	17.33
	3	16.89	16.94	17.00	17.08	17.16	17.27	17.39	17.53	17.71	17.91	18.15
	2	17.81	17.85	17.91	17.98	18.07	18.17	18.28	18.42	18.59	18.78	19.02
	1	18.74	18.79	18.85	18.92	19.00	19.09	19.21	19.34	19.50	19.69	19.92
	0	19.71	19.75	19.81	19.88	19.96	20.05	20.16	20.30	20.46	20.64	20.87
		0	1	2	3	4	5	6	7	8	9	10
		Initial number of unusable bicycles										

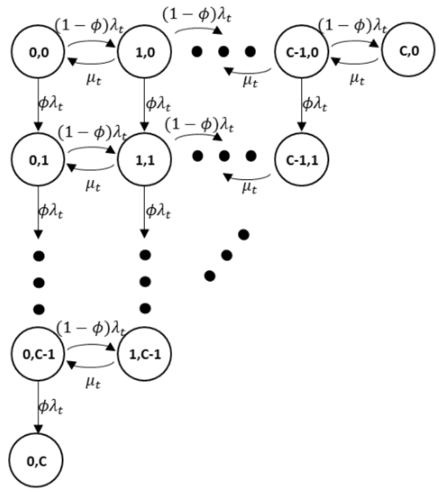

Figure 1 - Continuous-time Markov chain that represents the dynamics of the usable and unusable bicycle inventory levels

Appendix A1

Table 7: List of notations used throughout the paper, in order of appearance

Parameter	Definition
T	Length of the given period.
I_0	Initial number of usable bicycles in the station
B_0	Initial number of unusable bicycles in the station
p	Penalty for each bicycle shortage
h	Penalty for each locker shortage
C	Capacity of the station (number of lockers)
E^R	Time epochs in which the demands for bicycles or lockers occur under demand realization R
$I_j^R\left(I_0, B_0\right)$	The inventory level of usable bicycles right after the occurrence of the j^{th} demand in demand realization R, given the initial inventory of usable and unusable bicycles
$B_j^R\left(I_0, B_0\right)$	The inventory level of unusable bicycles right after the occurrence of the j^{th} demand in demand realization R, given the initial inventory of usable and unusable bicycles
$\left(d_j^{R,I}, d_j^{R,B}\right)$	The demand for bicycles or lockers at the j^{th} demand occurrence in demand realization R, $\left(d_j^{R,I}, d_j^{R,B}\right) \in \left\{(1,0),(-1,0),(0,-1)\right\}$
$\Delta_j^R\left(I_0, B_0\right)$	Bicycle shortage indicator right after the occurrence of the j^{th} demand in demand realization R, given the initial inventory of usable and unusable bicycles
$\Theta_j^R\left(I_0, B_0\right)$	Locker shortage indicator right after the occurrence of the j^{th} demand in demand realization R, given the initial inventory of usable and unusable bicycles
$F^R\left(I_0, B_0\right)$	User dissatisfaction under demand realization R, given the initial inventory of usable and unusable bicycles
$F\left(I_0, B_0\right)$	User dissatisfaction given the initial inventory of usable and unusable bicycles
Ω	A sequence of demand occurrences which do not include a returning of an unusable bicycle
$F^\Omega\left(I_0, B_0\right)$	User dissatisfaction under sequence of demand occurrences Ω, given the initial inventory of usable and unusable bicycles
ϕ	The probability that a bicycle is returned to a station in an unusable

	condition
θ	Range of the EUDF. The set of all possible inventory states $\{(I_0, B_0) \in \mathbb{Z}^2 \| I_0 \geq 0, B_0 \geq 0, I_0 + B_0 \leq C\}$
μ_t	Arrival rate of returners at time period t
λ_t	Arrival rate of renters at time period t

Appendix A2 – Unusable bicycles collection between repositioning visits

In Section 2, we introduced the assumption that during the period $[0,T]$, the inventory level is not externally altered. That is, until time T, no repositioning or repairing activities are performed in the station. An underlying assumption is that the collection of unusable bicycles is performed during some of the visits of repositioning trucks but not necessarily at all of them. We next demonstrate how the model can be adapted to cases in which the collection of unusable bicycles is performed between repositioning visits.

Recall that 0 and T represent the times of two consecutive visits of repositioning trucks. Now assume that a designated truck that collects unusable bicycles visits the station at time T' where $0 < T' < T$. We assume that if a designated collection truck is sent to the station, it collects all unusable bicycles in the station. However, it has no control over the inventory level of usable bicycles. If this visit is planned ahead, and is known at time 0, it may affect the replenishment decisions regarding usable bicycles at time 0, and therefore should be taken into account.

We adapt the EUDF presented in Equation (1) by separating the given period to two sub-periods $[0,T']$ and $[T',T]$ and refer to the EUDF in these two periods as F_1 and F_2 respectively.

Let $I_{T'}$ represent the inventory level of usable bicycles at time T', and recall that at time T' all unusable bicycles are removed from the station. We denote the adapted EUDF by $\tilde{F}(I_0, B_0)$ and it is calculated as follows:

$$\tilde{F}(I_0, B_0) \equiv F_1(I_0, B_0) + \sum_{I_{T'}=0}^{C} P_{(I_0, B_0), I_{T'}} \cdot F_2(I_{T'}, 0)$$

Where $P_{(I_0, B_0), I_{T'}}$ is the probability that at time T' there are $I_{T'}$ usable bicycles at the station, given that at time 0 there were I_0 usable bicycles and B_0 unusable bicycles in the station.

Note that the two instances of the EUDF, F_1 and F_2, are convex extensible under Assumption 1. The class of convex-extensible functions is closed under addition (Murota and Shioura, 2001). It is also easy to show that it is closed under multiplication by a positive scalar. Therefore, the resulting function is also convex-extensible under Assumption 1.

Appendix A3 – An example for a demand realization for which the EUDF is non-convex.

For a given demand realization R, showing that the function is convex in the initial inventory of unusable bicycles requires proving the following inequality:

$$F^R(I_0, B_0+2) - F^R(I_0, B_0+1) \geq F^R(I_0, B_0+1) - F^R(I_0, B_0)$$

Consider a small station with a capacity of two lockers in which the initial inventory of usable bicycles is zero. We focus on the following inequality:

$$F^R(0,2) - F^R(0,1) \geq F^R(0,1) - F^R(0,0)$$

We construct a demand realization with the following sequence of demand occurrences: The first demand occurrence is an attempt to return an unusable bicycle to the station. Then, we have a pair of demand occurrences, an attempt to return a usable bicycle followed by a rent attempt. This pair of demand occurrences is replicated n times (return unusable -> return usable -> rent -> return usable -> rent -> return usable -> rent ->…). In total, this demand realization consists of $2n+1$ demand occurrences. For this demand realization, the number of shortages is: $F^R(0,2)=2n+1$,

$F^R(0,1)=2n$ and $F^R(0,0)=0$.

Observing the inequality again, we obtain : $1>2n$. Note that n can be as large as desirable and therefore the inequality can be violated by any desired value.

Appendix A4 – Proof of Lemma 2

Lemma 2: under Assumption 1, the EUDF $F(I_0,B_0)$ is convex in the initial inventory of unusable bicycles B_0, i.e.: $F(I_0,B_0+2)-F(I_0,B_0+1)\geq F(I_0,B_0+1)-F(I_0,B_0)$

Proof: Consider the shortage occurrences given three settings of a station: (I_0,B_0+2), (I_0,B_0+1) and (I_0,B_0), i.e., when the number of usable bicycles is identical, but the number of unusable bicycles differs by one and two. We study the bicycle shortage indicator differences, $\Delta_j^R(I_0,B_0+2)-\Delta_j^R(I_0,B_0+1)$ and $\Delta_j^R(I_0,B_0+1)-\Delta_j^R(I_0,B_0)$ and demonstrate that each shortage occurrence in which the latter equals 1 is preceded by at least one shortage occurrence in which the former equals 1. This is also demonstrated for locker shortages. Therefore, by summing over all demand occurrences we obtain that the inequality holds for any demand realization that satisfies Assumption 1.

We distinguish between eight possible shortage events, as described in Table 8 and present the shortage differences in Table 9:

Table 8: Shortage events in three settings with initial inventories $\left(\boldsymbol{I}_0, \boldsymbol{B}_0+2\right)$ $\left(\boldsymbol{I}_0, \boldsymbol{B}_0+1\right)$ and $\left(\boldsymbol{I}_0, \boldsymbol{B}_0\right)$

Shortage event	Usable bicycles before the shortage occurs	Shortage type	Usable bicycles after the shortage occurs
1	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)-1=I_{j-1}^R\left(I_0,B_0\right)-2$	Bicycle	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)=I_j^R\left(I_0,B_0\right)-1$
2	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)-1=I_{j-1}^R\left(I_0,B_0\right)-1$	Bicycle	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)=I_j^R\left(I_0,B_0\right)$
3	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)=I_{j-1}^R\left(I_0,B_0\right)-1$	Bicycle	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)=I_j^R\left(I_0,B_0\right)$
4	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)=I_{j-1}^R\left(I_0,B_0\right)$	Bicycle	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)=I_j^R\left(I_0,B_0\right)$
5	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)-1=I_{j-1}^R\left(I_0,B_0\right)-2$	Locker	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)-1=I_j^R\left(I_0,B_0\right)-2$
6	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)-1=I_{j-1}^R\left(I_0,B_0\right)-1$	Locker	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)-1=I_j^R\left(I_0,B_0\right)-2$
7	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)=I_{j-1}^R\left(I_0,B_0\right)-1$	Locker	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)-1=I_j^R\left(I_0,B_0\right)-2$
8	$I_{j-1}^R\left(I_0,B_0+2\right)=I_{j-1}^R\left(I_0,B_0+1\right)=I_{j-1}^R\left(I_0,B_0\right)$	Locker	$I_j^R\left(I_0,B_0+2\right)=I_j^R\left(I_0,B_0+1\right)-1=I_j^R\left(I_0,B_0\right)-1$

Table 9: Shortage indicator differences

Shortage event	$\Delta_j^R\left(I_0,B_0+2\right)-\Delta_j^R\left(I_0,B_0+1\right)$	$\Delta_j^R\left(I_0,B_0+1\right)-\Delta_j^R\left(I_0,B_0\right)$
1	1	0
2	1	0
3	0	1
4	0	0
Shortage event	$\Theta_j^R\left(I_0,B_0+2\right)-\Theta_j^R\left(I_0,B_0+1\right)$	$\Theta_j^R\left(I_0,B_0+1\right)-\Theta_j^R\left(I_0,B_0\right)$
5	0	0
6	0	1

7	1	0
8	1	0

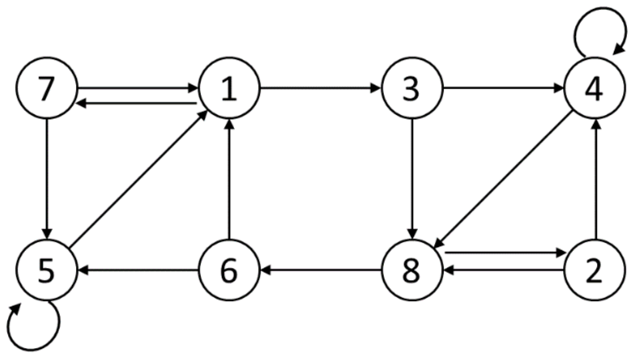

Figure 2: Possible transitions between shortage events

In Figure the possible transitions between shortage events are presented. Since at time $t=0$ (before the first demand occurrence) we have $I_0^R\left(I_0,B_0+2\right)=I_0^R\left(I_0,B_0+1\right)=I_0^R\left(I_0,B_0\right)=I_0$, the first shortage event may be either 4 or 8. Notice that each occurrence of event 6 is preceded by at least one occurrence of event 8. Similarly each occurrence of event 3 is preceded by at least one occurrence of event 1. Therefore by summing over all the demand occurrences of a given demand realization, we obtain:

$$\sum_{j=1}^{|E^R|}\left(\Delta_j^R\left(I_0,B_0+2\right)-\Delta_j^R\left(I_0,B_0+1\right)\right)\geq\sum_{j=1}^{|E^R|}\left(\Delta_j^R\left(I_0,B_0+1\right)-\Delta_j^R\left(I_0,B_0\right)\right)$$

and:

$$\sum_{j=1}^{|E^R|}\left(\Theta_j^R\left(I_0,B_0+2\right)-\Theta_j^R\left(I_0,B_0+1\right)\right)\geq\sum_{j=1}^{|E^R|}\left(\Theta_j^R\left(I_0,B_0+1\right)-\Theta_j^R\left(I_0,B_0\right)\right)$$

By multiplying the above inequalities by the relevant shortage penalties and summing the two inequalities we obtain:

$$F^R(I_0, B_0+2) - F^R(I_0, B_0+1) \geq F^R(I_0, B_0+1) - F^R(I_0, B_0).$$

Since this inequality holds for each demand realization that satisfies Assumption 1, it also holds for the expectation. �

Appendix A5

Table 10: Numerical results for Capital Bikeshare and the Citi Bike systems $(T = 3)$

Penalty Weights	Probability that a bicycle is returned unusable	Capital Bikeshare system			Citi Bike system		
		Number of stations with no gaps(out of 232)	Expected daily bicycle and locker shortages ($T = 1day$)		Number of stations with no gaps (out of 332)	Expected daily bicycle and locker shortages ($T = 1day$)	
			Maximal absolute gap	Maximal relative gap		Maximal absolute gap	Maximal relative gap
$p=1$ $h=1$	0	209	0.000288	0.000005	262	0.003545	0.000012
	0.01	212	0.000289	0.000005	269	0.002293	0.000010
	0.02	219	0.000291	0.000005	277	0.001602	0.000005
	0.03	218	0.000294	0.000005	268	0.001526	0.000006
	0.04	218	0.000311	0.000005	250	0.001525	0.000008
	0.05	217	0.000799	0.000005	218	0.001522	0.000008
	0.20	143	0.001634	0.000011	73	0.006039	0.000027
$p=0.75$ $h=1.25$	0	210	0.000214	0.000006	264	0.002486	0.000013
	0.01	212	0.000214	0.000004	268	0.001598	0.000011
	0.02	217	0.000216	0.000004	274	0.001157	0.000007
	0.03	218	0.000217	0.000004	268	0.001291	0.000005
	0.04	218	0.000443	0.000004	248	0.001239	0.000008
	0.05	218	0.001035	0.000006	220	0.001180	0.000010
	0.20	146	0.001699	0.000012	74	0.005275	0.000027
$p=0.5$ $h=1.5$	0	209	0.000173	0.000007	266	0.001442	0.000013
	0.01	211	0.000140	0.000004	267	0.000905	0.000012
	0.02	216	0.000141	0.000004	277	0.000726	0.000009
	0.03	218	0.000142	0.000004	271	0.001414	0.000005
	0.04	220	0.000576	0.000005	247	0.001306	0.000009
	0.05	218	0.001280	0.000008	220	0.001486	0.000012
	0.20	148	0.001764	0.000012	74	0.005857	0.000029

Appendix A6 – Bulk arrivals of renters and returners

In the approximation of the EUDF presented in Section 4, we assumed that the arrival processes of renters and returners to the station are time heterogeneous Poisson processes. In this appendix, we examine the accuracy of the polyhedral approximation under a more general arrival process. In particular, we consider arrival of users (renters and returners) in bulks of random sizes. The process that governs the arrival of bulks is a heterogeneous Poisson process. First, we demonstrate how the EUDF can be approximated under such an arrival process. Second, we estimate the distribution of bulks using the trip data from the Citi Bike case study. Third, using the estimated distribution of bulks, we reevaluate the EUDF values for all the stations in the two case studies and fit convex polyhedral functions as described in Section 4.

Recall that we denoted by μ_t and λ_t, the arrival rates of renters and returners, respectively. In addition, we denoted by ϕ the probability that when a user returns a bicycle to the station, the bicycle is unusable.

Let $a_{(I,B),(I',B')}$(t) denote the transition rate from state (I,B) to a different state (I',B') at time t. The transition rate matrix of the continuous-time Markov-chain represented in Figure 1 can be summarized as follows:

$$a_{(I,B),(I',B')}(t)=\begin{cases}(1-\phi)\lambda_t & if\ \ I'=I+1, B'=B, I+B+1<C\\ \phi\lambda_t & if\ I'=I, B'=B+1, I+B<C\\ \mu_t & if\ I'=I-1, B'=B, I>0\\ 0 & otherwise\end{cases}$$

and:

$$a_{(I,B),(I,B)}(t) = -\sum_{(I',B')\neq(I,B)} a_{(I,B),(I',B')}$$

Next, we adapt this continuous-time Markov chain to explicitly represent the bulk arrival of renters and returners. We assume that the arrival processes of bulks of renters and bulks of returners to the station are time heterogeneous Poisson processes, with rates $\tilde{\mu}_t$ and $\tilde{\lambda}_t$, respectively. The probability that the size of the arriving bulk is n is denoted by p_n. In addition, we denote by k the number of unusable bicycles returned to the station in a bulk of returners (naturally, $k \leq n$). The transition matrix of the resulting continuous-time Markov-chain can be summarized as follows:

and:

$$a_{(I,B),(I,B)}(t) = -\sum_{(I',B')\neq(I,B)} a_{(I,B),(I',B')}$$

To estimate the distribution of bulk sizes we take the following approach. Given trip history transactions data and given a time discretization, we consider all the demands of a certain type (renters or returners) at a station in each period of a given length as a bulk. This counting approach, neglects the in between arrivals of the other type of demand (returners or renters) and thus may over estimate the size of the bulks. Clearly, the longer the aggregation time unit is, we would exhibit larger bulks.

We have used the Citi Bike transactions data (July-August 2014, weekdays), and tested two time discretizations, namely, 1 minute and 2 minutes. During that period there were in total 1,579,304 transactions. In Table 11, we present the distribution of bulk sizes of renters and returners according to 1 minute and 2 minute aggregations.

Table 11: Bulk distribution in the Citi Bike data according to 1 minute and 2 minutes aggregations

Bulk Size	1-minute aggregation				2-minute aggregation			
	Rent		Return		Rent		Return	
	Count	Pct.	Count	Pct.	Count	Pct.	Count	Pct.
1	1,194,311	87.147%	1,204,210	87.401%	981,591	79.241%	996,980	79.618%
2	150,160	10.957%	150,080	10.893%	198,462	16.021%	200,063	15.977%
3	21,114	1.541%	19,876	1.443%	42,852	3.459%	42,480	3.392%
4	3,650	0.266%	2,995	0.217%	10,814	0.873%	9,657	0.771%
5	838	0.061%	523	0.038%	3,032	0.245%	2,280	0.182%
6 and above	382	0.028%	113	0.008%	1,990	0.161%	744	0.059%

Based on the 2-minute aggregation, we approximated the EUDF for the two studied systems using the following distribution of bulk sizes for both renters and returners: $p_1 = 80\%, p_2 = 16\%, p_3 = 3\%, p_4 = 0.8\%, p_5 = 0.2\%.$ Note that in stations that exhibit high demand rates, it is likely that more than one bulk will arrive in a time unit. In such cases, the arrival of more than 5 renters/returners in a time unit occurs with a probability that is not negligible. The arrival rates of bulks were obtained by dividing all the arrival rates of renters and returners by the expected bulk size (in this case 1.252). In Table 12 we present the numerical results for a time horizon of one day $(T = 1)$. It is observable that the gaps presented in this table are significantly smaller than those presented in Table 5. That is, taking into account explicitly the arrival of bulks results with a more accurate approximation of the EUDF by a convex polyhedral function. The reason for this is that the arrivals of bulks have increasing effect when the station in its initial state is closer to being empty or full which "contributes" to the convexity of the functions.

Table 12: Numerical results for Capital Bikeshare and the Citi Bike systems (Bulks, $T = 1$)

Penalty Weights	Probability that a bicycle is returned unusable	Capital Bikeshare system			Citi Bike system		
		Number of stations with no gaps (out of 232)	Expected daily bicycle and locker shortages ($T = 1day$)		Number of stations with no gaps (out of 332)	Expected daily bicycle and locker shortages ($T = 1day$)	
			Maximal absolute gap	Maximal relative gap		Maximal absolute gap	Maximal relative gap
$p = 1$ $h = 1$	0	227	0.000020	0.000000	307	0.000442	0.000063
	0.01	227	0.000022	0.000000	307	0.000439	0.000049
	0.02	226	0.000024	0.000000	307	0.000436	0.000043
	0.03	228	0.000026	0.000000	308	0.000432	0.000021
	0.04	228	0.000027	0.000000	308	0.000430	0.000004
	0.05	227	0.000029	0.000000	310	0.000426	0.000004
	0.20	222	0.000625	0.000006	213	0.002747	0.000057
$p = 0.75$ $h = 1.25$	0	227	0.000013	0.000000	305	0.000398	0.000057
	0.01	227	0.000012	0.000000	307	0.000380	0.000064
	0.02	226	0.000012	0.000000	308	0.000324	0.000054
	0.03	227	0.000014	0.000000	308	0.000300	0.000033
	0.04	224	0.000014	0.000000	309	0.000298	0.000023
	0.05	226	0.000017	0.000000	308	0.000294	0.000003
	0.20	221	0.000754	0.000007	212	0.002823	0.000050
$p = 0.5$ $h = 1.5$	0	225	0.000018	0.000000	307	0.000479	0.000058
	0.01	228	0.000018	0.000000	307	0.000467	0.000057
	0.02	228	0.000015	0.000000	306	0.000428	0.000051
	0.03	228	0.000012	0.000000	307	0.000365	0.000041
	0.04	228	0.000010	0.000000	308	0.000276	0.000033
	0.05	228	0.000006	0.000000	307	0.000188	0.000022
	0.20	224	0.000891	0.000008	214	0.002900	0.000046

Mor Kaspi is a Postdoctoral researcher in the Urban Transport Systems Laboratory (LUTS) at École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. He holds a B.Sc. (2008), a M.Sc. (2010) and a Ph.D. (2016) in Industrial Engineering from the Iby and Aladar Fleischman Faculty of Engineering, Tel Aviv University, Israel. His research interests are in operations management of transportation and logistics systems, focusing on modeling and optimization techniques.

Tal Raviv is a member in the Industrial Engineering department at the Iby and Aladar Fleischman Faculty of Engineering, Tel Aviv University, Israel. He holds a BA from the Eitan Berglas

School of Economics, Tel Aviv University (1993), an MBA from the Recanati School of Business, Tel Aviv University (1997), and a PhD in Operations Research from the William Davidson Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa (2003). He spent two years (2004-2006) as a postdoctoral fellow in the Sauder School of Business at University of British Columbia, Vancouver, Canada. His current primary research interest is in transportation and logistics with a focus on sustainable transportation and shared mobility systems.

Michal Tzur is a professor in the Industrial Engineering department at the Faculty of Engineering, Tel Aviv University, Israel. She received her Ph.D. in Management Science from the Graduate School of Business at Columbia University in the city of New York. She was a faculty member in the business school of the University of Pennsylvania (Wharton School) and a visiting faculty member in the department of Industrial Engineering and Management Sciences (IEMS) at Northwestern University. Michal's main research interests are in the area Humanitarian Logistics, Vehicle Sharing Systems, Transportation Systems and Supply Chain Management. Since 2015 she is the president of the Operations Research Society of Israel (ORSIS).