A Traveler Incentive Program for Promoting Community-Based Ridesharing

Amirmahdi Tafreshian[a], Neda Masoud[a]
[a] Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan 48109
Contact: atafresh@umich.edu (AT); nmasoud@umich.edu, https://orcid.org/0000-0002-6526-3317 (NM)

Abstract. Traffic congestion has become a serious issue around the globe, partly owing to single-occupancy commuter trips. Ridesharing can present a suitable alternative for serving commuter trips. However, there are several important obstacles that impede ridesharing systems from becoming a viable mode of transportation, including the lack of a guarantee for a ride back home as well as the difficulty of obtaining a critical mass of participants. This paper addresses these obstacles by introducing a Traveler Incentive Program (TIP) to promote community-based ridesharing with a ride-back home guarantee among commuters. The TIP program allocates incentives to (1) directly subsidize a select set of ridesharing rides, and (2) encourage a few, carefully selected set of travelers to change their travel behavior (i.e., departure or arrival times). We formulate the underlying ride-matching problem as a budget-constrained min-cost flow problem, and present a Lagrangian Relaxation-based algorithm with a worst-case optimality bound to solve large-scale instances of this problem in polynomial time. We further propose a polynomial-time budget-balanced version of the problem. Numerical experiments suggest that allocating subsidies to change travel behavior is significantly more beneficial than directly subsidizing rides. Furthermore, using a flat tax rate as low as 1% can double the system's social welfare in the budget-balanced variant of the incentive program.

Keywords: P2P ridesharing; Incentive design; Community-based ridesharing; Monetary subsidy; Budget-constrained min-cost flow problem; Guaranteed ride-back home

1 Introduction

In recent years, traffic congestion has become a serious issue around the globe. In a contemporary study of 1,360 cities in 38 countries, Bloomberg CITYLab asserted that traffic congestion costs over $305 Billion per year only in the U.S. (Schneider 2018). This spike in traffic congestion, especially during morning and evening peak hours, is mainly due to the rising number of solo-driver commuting trips. Despite tremendous expenditure on subsidy, public transit services have failed to alleviate congestion by shifting solo driving toward more sustainable forms of transport. In addition, not only has the recent proliferation of Transportation Network Companies (TNCs) such as Uber and Lyft not addressed this issue, but it has exacerbated congestion, particularly in large cities (Hawkins 2019).

Peer-to-peer (P2P) ridesharing is a manifestation of the sharing economy business model in the mobility market, and provides a promising solution for mitigating traffic congestion. In contrast to TNCs that produce high empty miles, thereby adversely affecting traffic congestion, P2P ridesharing improves congestion by increasing the utilization rate of empty seats. Additionally, P2P ridesharing provides a unique opportunity for communities to augment transit services by serving their mobility needs internally. In community-based ridesharing — a form of P2P ridesharing for commuters — the members of the community who own cars, henceforth referred to as drivers, transport their peers, henceforth referred to as riders, along their routes while completing their own personal trips. Aside from its environmental benefits, community-based ridesharing provides a promising mobility solution for the following reasons:

1. **Existing trust**: Drivers and riders are members of the same small community, preventing lack of participation that may arise from lack of trust in large metropolitan areas.
2. **Lack of opportunity cost**: Drivers travel according to their own schedules, and may only take small detours to serve their fellow community members.
3. **Revenue for the community**: The fare of the rides will be set so as to compensate drivers fully for their detours and partially for their base travel costs, as they would be sharing the cost of their base trips with riders.

Despite abundant benefits of ridesharing systems, there are a few obstacles that hinder the adoption of such systems by commuters in practice. First, commuters are in general reluctant to leave their vehicles at home in favor of outsourcing their rides in the morning if they do not have a guarantee for a ride back home. Traditional carpooling services can provide such a guarantee for commuters who have fixed and common working hours. However, carpooling may not be a viable option for commuters whose working hours may shift from one day to the next. Secondly, in P2P ridesharing participants are available in the network only for a short period of time (compared to TNC drivers and transit providers), which leads to low spatio-temporal proximity among trips. These factors pose a challenge for P2P ridesharing systems achieving a critical mass, and thereby, becoming a viable transportation option in the long run.

This paper tackles this timely issue by designing a traveler incentive program (TIP) that allocates monetary subsidies to riders and drivers to foster ridesharing participation rates and provide an opportunity for the commuters within a community to serve their mobility needs internally. To this end, this paper develops a ridesharing system with ride-back guarantee that incentivizes commuters to share their morning and evening trips with their peers. TIP finds the optimal allocation of a set budget to participants so as to: (i) subsidize a selective set of rides, and (ii) change the travel behavior of a small, carefully selected set of commuters, with respect to their travel time windows. The ultimate goal of TIP is to maximize social welfare of system participants while ensuring that every dollar injected to the system will generate a higher value in social welfare.

In the rest of this paper, Section 2 provides a review of the literature related to P2P ridesharing and using incentives to promote shared mobility services. Next, we carefully define the problem in this paper and its underlying assumptions in Section 3. In Section 4, we present a mathematical formulation for the problem and a solution methodology to solve its large instances. Section 5 presents the results of several numerical experiments that evaluate different aspects of our proposed methodology. Finally, Section 6 finalizes this paper by summarizing our findings and providing directions for future research.

2 Literature Review

In this section, we first present a brief overview of the literature in P2P ridesharing, followed by a review of studies in shared mobility that consider various types of incentives to promote their systems. Finally, we clearly state the contributions of this paper.

2.1 Peer-to-Peer (P2P) Ridesharing

P2P ridesharing is a shared mobility platform that encourages users with similar routes and time schedules to share their rides together (Agatz et al. 2012). In spite of some similarities, one must distinguish P2P ridesharing from other forms of shared mobility platforms such as carpooling (see e.g., Baldacci, Maniezzo, and Mingozzi (2004)), since it does not require long-time commitments from users, as well as ride-sourcing (see e.g., Xu, Yin, and Ye (2020)) or taxi-sharing (see e.g., Alonso-Mora et al. (2017)) since drivers in P2P ridesharing are not treated as employees and their primary intention is to complete their own personal trips. In what follows, we describe a number of major characteristics of P2P ridesharing, henceforth referred to as ridesharing, and its variant forms. For further information on ridesharing, the interested reader is referred to the surveys by Agatz et al. (2012; Furuhata et al. 2013; Tafreshian, Masoud, and Yin 2020).

Ridesharing systems can be roughly divided into the two categories of static ridesharing (see e.g., Regue, Masoud, and Recker (2016; Xiaoqing Wang, Dessouky, and Ordonez 2016; Long et al. 2018)) and dynamic ridesharing (see e.g., Agatz et al. (2011; Lee and Savelsbergh 2015; Masoud and Jayakrishnan 2017b)). In the former case, the trip information of all users is known ahead of time, while in the latter case, users enter the system dynamically and register their trips shortly before their departure times. Most of the studies in P2P ridesharing consider the sets of riders and drivers as two mutually exclusive sets (see e.g., Agatz et al. (2011; Nourinejad and Roorda 2016; Najmi, Rey, and Rashidi 2017)). However, a few studies relax this assumption and let the system operator decide the most beneficial role (i.e., rider or driver) for each participant (see e.g., Amey (2011; Chen et al. 2019; Tafreshian and Masoud 2020a, 2020b)). The core of a ridesharing system is a ride-matching problem, the solution of which determines the optimal assignment between riders and drivers, users' trip schedules, and drivers' routes. With an intent to make ridesharing convenient for both riders and drivers, many studies consider the simplest form of the ride-matching problem in which every user can be matched with at most one other user (see e.g., Ma, Zheng, and Wolfson (2013; Najmi, Rey, and Rashidi 2017; Xing Wang, Agatz, and Erera 2017)). In order to increase the possibility of matching, however, a number of studies diverge from this assumption by allowing multiple riders per vehicle and transfers between vehicles (see e.g., Stiglic et al. (2015; Masoud and Jayakrishnan 2017a; Chen et al. 2019)). Moreover, a number of studies incorporate the choice of ride-back home guarantee in their models, which motivates rider participation, and increases the level-of-service offered by the system (see e.g., Regue, Masoud, and Recker (2016; Lloret-Batlle, Masoud, and Nam 2017; Chen et al. 2019; Hasan, Van Hentenryck, and Legrain 2020)).

Based on the definitions above, the community-based ridesharing proposed in this paper can be categorized as a static one-to-one ridesharing system with ride-back guarantee. The choice of a static system is supported by the outcome of a ridesharing survey in Berkeley, CA, which concludes that commuters prefer to learn about their ride-share arrangements at least a night before (Deakin, Frick, and Shively 2010). Also, given the assumption that all users select the shortest travel time path as their selected route, one-to-one ride-matching ensures that (i) riders do not experience any detour during their travel, and (ii) drivers' inconvenience due to pick-ups and drop-offs are minimized. Finally, the results of a behavioral study by Brownstone and Golob (1992) indicates that the option of ride-back home guarantee motivates a high percentage of commuters to engage in ridesharing programs.

2.2 Incentives in Shared Mobility

Since the introduction of shared mobility services, several studies have emphasized the need for designing incentives to motivate solo drivers to participate in ride-share programs. The proposed incentives can roughly fall into two categories of indirect and direct (a.k.a. financial subsidies) incentives. The employer-provided parking discounts for high occupancy vehicles (HOV) is an example of an indirect incentive that can significantly incentivize daily commuters to shift toward ridesharing (Brownstone and Golob 1992; Su and Zhou 2012). Another indirect incentive that proves useful in practice is the possibility of using freeway HOV lanes that leads to savings in commute times (Brownstone and Golob 1992; Lloret-Batlle, Masoud, and Nam 2017). Finally, integrating ridesharing with public transit and/or other transport programs can reduce the travel cost of commuting and encourage a higher number of commuters to share their rides (Deakin, Frick, and Shively 2010; Nam et al. 2018; Bian and Liu 2019).

Aside from indirect incentives, many studies stress the necessity of adopting various form of financial subsidiary schemes to promote ridesharing among commuters (Chan and Shaheen 2012; Agatz et al. 2012). The ultimate goal of financial (or direct) incentives is to maximize fleet utilization, and thereby reduce traffic congestion. Stiglic et al. (2016) conduct a comprehensive case study to evaluate the effect of time flexibility on the matching rate in one-to-one ridesharing systems. Based on the results of their experiments, they emphasize the importance of adopting an incentive scheme that provides monetary benefits to commuters to increase their time flexibility. In a study that focuses on the role of rider-driver cost-sharing strategies in the success of ridesharing programs, Xiaolei Wang, Yang, and Zhu (2018) show that providing ridesharing users with sufficient subsidies can reduce the cost of participation and turn ridesharing into a viable alternative to public transit. They further emphasize the need for designing appropriate subsidizing schemes.

The consideration of financial subsidies is not limited to ridesharing systems and has been studied in other types of shared mobility services. Qian et al. (2017), for instance, introduce ride incentives for groups of passengers in a shared-taxi service as discounts towards their trip fares. They further propose different algorithms that find the best ride incentives to improve total saved mileage. For the operation of ride-sourcing platforms such as Uber and DiDi Chuxing, Zhao and Chen (2019) compare the ex-ante and ex-post destination information models and show the effectiveness of subsidies in attracting more participants under the latter model. They further design a subsidy scheme based on the income of drivers that motivates them to serve farther distant passengers. Luo et al. (2019) develop a new dynamic games approach to find the optimal subsidy policy that accelerates the adoption of automated vehicles (AV's). In their approach, adaptive subsidies are computed based on the state of the AV market penetration process under uncertainty. They claim that the optimal subsidies further incentivize the AV manufacturers to improve their technology, and offer pricing incentives to potential consumers.

More recently, through a joint simulation of car-sharing, bike-sharing and ride-hailing for a city-scale transport system, Becker et al. (2020) found that the highest system-level impacts can be achieved when the operations of shared modes are subsidized. More interestingly, they showed that the total amount of subsidies required for these shared modes is lower than the amount paid for current regular public transport services. Finally, Xiong et al. (2020) propose an integrated system that provides personalized travel alternatives and monetary incentives for travelers. As a part of their system, to reduce traffic congestion, they offer alternative departure times for commuters and compensate them with monetary subsidies. This incentive scheme has been commercialized as a mobile app, called incenTrip, and is currently used in the Washington area. Finally, Li, Nie, and Liu (2020) considered a ridesharing system in which riders and drivers bid on their sensitivity to schedule displacement, and introduce a VCG mechanism which is individually rational, truthful, and both computationally and economically efficient. They further introduce a one-sided reward pricing policy that sacrifices truthfulness to satisfy weakly budget-balancedness.

Based on the findings of these studies, we propose two types of monetary incentives that can help improve the social welfare of a community adopting a ridesharing system. The first incentive is targeted toward changing travelers' schedules to increase their chance of being matched. The other incentive attempts to compensate the negative externalities of riders and drivers sharing rides together.

2.3 Our Contributions

The contributions of this paper are as follows:

- We introduce a traveler incentive program for a ridesharing system with guaranteed ride-back, and present a mixed integer nonlinear optimization model that determines the optimal matching, scheduling, and incentive allocation.
- We decompose the mixed integer nonlinear optimization model into a linear model that determines the optimal amount of incentives to each rider-driver pair, and a budget-constrained min-cost flow problem, which is known to be NP-complete.
- We propose a polynomial-time Lagrangian Relaxation method to efficiently find near-optimal solutions for large-scale instances of the problem, and provide a worst-case optimality bound for its performance.
- We propose a budget-balanced variant of the incentive program, which could be solved in polynomial time.
- We perform a comprehensive set of numerical experiments to showcase the impact of the proposed incentive program on serving the mobility needs in a city.

Comparing to the existing literature, this study is among the first to introduce an incentive program that promotes community-based ridesharing among daily commuters. We present the mathematical formulation for this problem, and devise a polynomial time methodology based on Lagrangian Relaxation to solve it. Also, to the best of our knowledge, we are the first to present a bounded solution for the budgeted min-cost max flow problem as a part of our proposed methodology.

3 Problem Statement

The main focus of this paper is a static one-to-one ridesharing system for commuters (although In section 5 we will investigate a dynamic setting). This program guarantees ride-back services for the riders who register both their morning and evening trips in the system. Let N denote the set of all participants (users) that register their trips in a given day. A participant in the system can be either a driver or a rider in a given day. Therefore, set N can be further partitioned into two disjoint sets of riders, denoted by R, and drivers, denoted by D. We assume that the ridesharing system knows the following information regarding each user $n \in N$:

- $\mathsf{F}(n)$: the value of every unit of time spent on traveling
- $\mathsf{H}(n)$: the value of every unit of distance driven

These values can either be specified directly by having the users answer a short survey, or estimated based on the provided information upon registration for the first time in the system (e.g., occupation, place of residence, vehicle's make and model, etc.). Every user $n \in N$ may register a trip in the morning, represented by n', a trip in the evening, represented by n'', or both. Let $N' = R' \cup D'$ and $N'' = R'' \cup D''$ respectively denote the sets of all trips in the morning and evening. Also, let $D''' \subset D$, $R''' \subset R$, $N''' = R''' \cup D'''$ denote the set of drivers, riders, and users who register both their morning and evening trips in the system, respectively. Due to the similarities between the characteristics of morning and evening trips, we describe our assumptions using only the morning trips in the rest of this section. Table A.1 summarizes the notation used in this paper.

It is assumed that the study region consists of a large number of stations from/at which trips originate/end. Thus, every trip $n' \in N'$ can be characterized by the following information:

- $\mathsf{I}(n')$: the origin station for the morning trip of user n
- $\mathsf{J}(n')$: the destination station for the morning trip of user n
- $\mathsf{T}(n')$: the desired earliest departure time of user n from the origin station in the morning
- $\mathsf{Q}(n')$: the desired latest arrival time of user n from the destination station in the morning

We further assume that the shortest-path travel time and driving distance between every pair of stations during the morning period are known and stored in the hash tables τ and ρ, respectively. Thus, $\tau_{i,j}/\rho_{i,j}$ represent the shortest-path travel time/distance from station i to station j. We refer to $[\mathsf{T}(n'), \mathsf{Q}(n')]$ as the morning time window of user n. The length of this time window is rather tight, but always greater than or equal to $\tau_{\mathsf{I}(n'),\mathsf{J}(n')}$.

To provide a high quality of service for both riders and drivers participating in the ridesharing system, we assume that in a given time period (e.g., morning), each driver will give a ride to at most one rider, and riders complete their trips with at most one driver. This limits the length of detours incurred by drivers, and guarantees no detour and transfer for riders.

3.1 Static One-to-One Ridesharing with Ride-Back Guarantee

The problem of one-to-one ridematching with guaranteed ride-backs was first considered by Agatz et al. (2011), where they formulated it as a weighted matching problem with the addition of a set of bundle constraints that relate riders' inbound trips to their outbound trips. In general, this problem may no longer have the unimodularity property, and thus, requires an MIP solver to solve. However, in the cases where inbound and outbound trips occur in two non-overlapping periods (e.g., all the outbound trips occur in the morning peak hours and inbound trips occur in the evening peak hours), Lloret-Batlle, Masoud, and Nam (2017) show that the ridematching problem can be formulated as a min-cost max flow problem in a weighted directed graph $G = (V, E)$, where the set of nodes is denoted by $V = \{\mathsf{s}, \mathsf{t}\} \cup N' \cup N''$, and the edge set is denoted by E. Figure 1 shows an example of such a graph with 11 users. The auxiliary nodes s and t respectively represent the source node and the target node. The edge set E consists of two edge subsets: (i) the auxiliary edges with capacity of 1 unit and 0 units of cost, and (ii) the potential matching edges with capacity of 1 unit and $-\mathsf{W}$ units of cost. The auxiliary edges connect (i) node s to all driver trips $d' \in D'$ and all rider trips $r \in R''$ when $r \notin R'''$, (ii) all driver trips $d'' \in D''$ and all rider trips $r' \in R'$ when $r \notin R'''$ to target node t, and (iii) every rider trip $r' \in R'$ to rider trip $r'' \in R''$ when $r \in R'''$. The potential matching edges connect the morning drivers to the morning riders and the evening riders to the evening drivers. A potential match edge exists in graph G if (i) the match is spatio-temporally feasible, and (ii) both parties prefer the match to their other available options outside the ridesharing system, i.e., the match is individually rational.

In what follows, we develop the mathematical equations for spatio-temporal feasibility and individual rationality conditions for the morning period. Similar equations can be derived for the evening period. Spatio-temporal feasibility requires a driver to be capable of providing a ride to a rider within the rider's time window, while completing their own trip within their time window. Thus, the following two equations must be satisfied simultaneously:

$$\max\{\mathsf{T}(d') + \tau_{\mathsf{I}(d'),\mathsf{I}(r')}, \mathsf{T}(r')\} + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} \leq \mathsf{Q}(r'), \tag{1a}$$

$$\max\{\mathsf{T}(d') + \tau_{\mathsf{I}(d'),\mathsf{I}(r')}, \mathsf{T}(r')\} + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} + \tau_{\mathsf{J}(r'),\mathsf{J}(d')} \leq \mathsf{Q}(d'). \tag{1b}$$

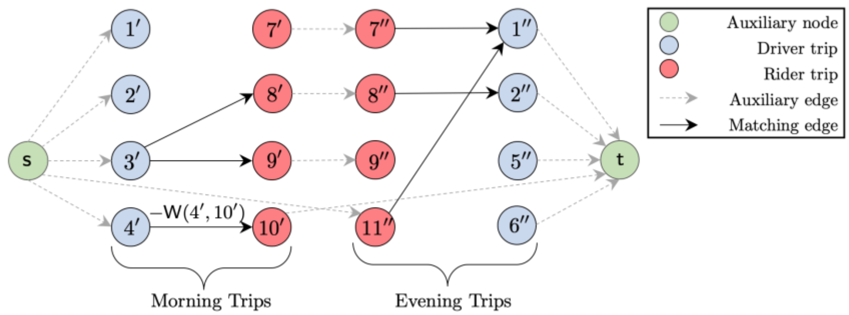


Figure 1. Graph $G = (V, E)$ for a ridesharing system with 11 users where $D' = \{1', 2', 3', 4'\}$, $D'' = \{1'', 2'', 5'', 6''\}$, $R' = \{7', 8', 9', 10'\}$, and $R'' = \{7'', 8'', 9'', 11''\}$.

Equations (1a) and (1b) respectively ensure that the match allows rider r and driver d to complete their trips within their specified time windows. While these equations describe the spatio-temporal feasibility of a match, for a match to be deemed valuable and accepted by both parties, the fare for the ride should be set to an amount that is acceptable by both the driver and the rider. Let $\mathsf{U}(r'|d')$ and $\mathsf{U}(d'|r')$ respectively denote the valuations of rider r and driver d of sharing a ride together. This valuation can be defined as the difference between the cost of sharing the ride and the cost of driving alone (in case the rider does not own a car, this cost is set to the cost of taking a taxi cab). Based on the elicited information from both parties, these valuations can be found as:

$$\begin{aligned} \mathsf{U}(r'|d') = \quad & \mathsf{H}(r)\, \rho_{\mathsf{I}(r'),\mathsf{J}(r')}\,, & \text{(2a)} \\ \mathsf{U}(d'|r') = \quad & -\mathsf{H}(d)\left(\rho_{\mathsf{I}(d'),\mathsf{I}(r')} + \rho_{\mathsf{I}(r'),\mathsf{J}(r')} + \rho_{\mathsf{J}(r'),\mathsf{J}(d')} - \rho_{\mathsf{I}(d'),\mathsf{J}(d')}\right) & \text{(2b)} \\ & -\mathsf{F}(d)\left(\tau_{\mathsf{I}(d'),\mathsf{I}(r')} + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} + \tau_{\mathsf{J}(r'),\mathsf{J}(d')} - \tau_{\mathsf{I}(d'),\mathsf{J}(d')}\right), \end{aligned}$$

where $\mathsf{F}(n) > 0$ and $\mathsf{H}(n) > 0$ respectively denote the values of time (in \$ per unit of time) and distance (in \$ per unit of distance) for user $n \in N$. The statements in parentheses in Equations (2b) respectively represent the distance and time of the detour incurred by driver d to serve rider r in the morning. If we assume that all users follow a quasi-linear utility, then the sum of valuations in (2a) and (2b) yields the potential monetary saving (gain) due to driver d providing a ride to rider r, denoted as $\mathsf{W}(d', r') = \mathsf{U}(d'|r') + \mathsf{U}(r'|d')$. If the gain for a match is non-negative, that is, if $\mathsf{W}(d', r') > 0$, then there exists a pricing mechanism to split the benefits between users r and d so as to ensure they both have non-negative utilities, i.e., both choices are individually rational (e.g., they share the profit based on the length of the two trips). Note that $\mathsf{W}(r'', d'') = \mathsf{U}(r''|d'') + \mathsf{U}(d''|r'')$ can be computed similarly for the evening trips.

3.2 Incentive Design

The graph corresponding to the ridesharing problem with guaranteed ride-back can be sparse, as demonstrated in the example in Figure 1. This sparsity is partly due to individuals' rather tight travel time windows, and partly a result of the heterogeneity in their valuations of options. It is not surprising to see trip requests with tight time windows, as this ensures the rides to be in congruence with travelers' preferences. Furthermore, it is realistic to assume that individuals who own cars are in a superior financial status than those who do not. As such, having drivers with higher values of time, and possibly distance, get compensated for their detours by riders may lead to the ridesharing option not being affordable for a large portion of riders. To tackle these issues, we introduce two types of incentives to increase the number of edges between trip nodes in graph G, eventually leading to higher percentage of matches, and possibly higher levels of social welfare in the community.

3.2.1 Behavioral Adjustment (BA) Incentives.

When registering a trip, the desired earliest departure time and latest arrival time are two of the trip characteristics that a participant has to specify. A wider time window for a driver implies a potentially longer detour and more flexibility in departure time. For riders, a wider time window only implies higher flexibility in departure time, as a rider's travel duration would be that of their shortest-path travel time. Clearly, by providing tighter time windows, participants would receive matches that are more congruent with their preferences. A behavioral adjustment (BA) incentive encourages participants to be more flexible with their travel time windows. Let us subsidize user n by an incentive rate of $\mathsf{F}(n)$, equal to their value of time, to widen their morning trip time window for $\gamma(n')$ units, with the assumption that their disutility from leaving earlier than $\mathsf{T}(n')$ and arriving later than $\mathsf{Q}(n')$ are the same and are proportional to $\mathsf{F}(n)$. If we let $\gamma^{-}(n')$ and $\gamma^{+}(n')$ respectively denote the amount of time extension in time window of trip n' from left and right, the adjusted time window can be shown as $[\mathsf{T}(n') - \gamma^{-}(n'), \mathsf{Q}(n') + \gamma^{+}(n')]$ subject to the constraint $\gamma^{-}(n') + \gamma^{+}(n') = \gamma(n')$. We further assume that the total extension in a trip time window cannot exceed a pre-defined threshold, denoted by Γ, which can be defined as the maximum time window extension offered to a participant, and can be either specified by the user as a part of their trip information during trip registration or set by the system operator based on the targeted level-of-service (LoS). Note that the right value for $\gamma(n')$ depends on other users' trip information; hence, it must be determined by the system operator as a part of the ridematching problem.

3.2.2 Individual Rationality (IR) Incentives.

Consider a rider-driver pair (d', r') for whom spatio-temporal conditions (i.e., Equations (1a) and (1b)) are satisfied, but the individual rationality conditions are violated (i.e., $\mathsf{W}(d', r') < 0$). An individual rationality (IR) incentive, denoted by $\lambda(d', r')$, provides subsidies that allow for introducing such a link in graph G with a non-negative gain.

The consideration of these incentives in our ridesharing system has two important consequences. First, the Equations in (1) can no longer be used to determine the spatio-temporal feasibility of a match, since the adjusted time windows depend on the unknown values of the BA incentives. Secondly, the rider and driver valuations from sharing a ride in (2) as well as the gains of potential matches are dependent upon the unknown values of both incentive types. These consequences clearly suggest that no longer can we find the optimal ridematching using the min-cost max flow problem described above. As such, in this paper we develop a ridematching problem that determines the optimal matching, trip scheduling, and incentive allocation that maximizes the social welfare given a monetary budget of B dollars. In order to make sure that the available budget is used wisely, we further require that every dollar spent on subsidy contributes more than one dollar to the system's social welfare.

4 Solution Methodology

In this section, we first present the mathematical formulation of the TIP for a simpler system involving only one-time trips with no ride-back guarantee (e.g., the morning trips) given the assumptions provided in the previous section. Next, we modify this mathematical formulation to model the TIP for a system that includes both the morning and evening trips. We propose an efficient algorithm to solve large-scale instances of this problem in a timely manner, and prove a worst-case optimality bound for its performance. Finally, we propose a budget-balanced counterpart of the TIP through taxation.

4.1 Mathematical Formulation For the Morning Trips

Let us consider a ridesharing system that includes only the morning trips N'. Also, let A' denote the set of all driver-rider trip pairs in the morning, i.e., $A' = \{(d', r') \in D' \times R'\}$. Based on the assumptions provided in Section 3, the optimal matching, scheduling, and incentive allocation of the system can be obtained by solving the mixed integer nonlinear program (MINLP) presented in (3). In this formulation, there are eight sets of decision variables. The decision variable $x(a')$ is a binary variable that holds the value 1 if the pair of trips in $a' \in A'$ share their rides together, and the value 0 otherwise. The continuous decision variables $t(n')$ and $q(n')$ denote the start and end time of the trip $n' \in N'$, respectively. The decision variable $\gamma(n')$ can be defined as the amount of extension in the time window of trip n'. We further define two variables, $\gamma^{-}(n')$ and $\gamma^{+}(n')$, to denote the amount of extension in the time window of the morning trip n' from the left and the right, respectively. Finally, the decision variables $\lambda(a')$ and $w(a')$ respectively represent the IR incentive and the gain for the match between a driver-rider trip pair $a' \in A'$.

Let ϵ be an infinitesimal positive value. The objective in (3a) maximizes the difference between the system's social welfare and the total amount of subsidies spent on the BA and IR incentives. Choosing this objective enables us to maximize social welfare while allocating subsidies only when the added value to social welfare is strictly higher than the amount of subsidy; that is, for each dollar spent on subsidy, a return-to-investment of more than one dollar can be obtained in social welfare. Constraint (3b) defines the adjusted gain of match (d', r') in the morning as the sum of the original savings due to driver d sharing their ride with rider r and the subsidies allocated to these participants and their coalition. Constraint (3c) ensures that if users r and d are determined to share a ride in the morning, i.e., if $x(d', r') = 1$, then their coalition is individually rational, i.e., $w(d', r') \geq 0$. Constraints (3d) and (3e) together guarantee that every trip starts and ends within its adjusted time windows. Constraint (3f) states that if driver trip d' is matched with rider trip r', the difference between their trips' start times must be at least equal to the shortest-path travel time between their origin stations. Constraint (3g) defines the end time of rider trip r' as the sum of its start time and the duration of the trip. Constraint (3h) ensures that the difference between the end time of driver trip d' and rider trip r', if matched together, must be equal to the shortest-path travel time from the rider's destination station to driver's destination station. Constraint (3i) defines the total extension in a trip's time window as the sum of the extensions from the left and the right. Constraint (3j) ensures that the time window extension cannot exceed its maximum limit. Constraints (3k) and (3l) respectively ensure that every rider is matched with at most one driver, and each driver serves at most one rider in the morning. Finally, Constraint (3m) ensures that the total amount of allocated subsidy is lower than the available budget. Constraints (3n)-(3p) are the non-negativity and integrality constraints.

$$\max \quad \sum_{a' \in A'} w(a')\, x(a') - (1+\epsilon)\Big(\sum_{n' \in N'} \mathsf{F}(n)\, \gamma(n') + \sum_{a' \in A'} \lambda(a')\Big) \tag{3a}$$

$$\text{s.t.} \quad w(d', r') = \mathsf{W}(d', r') + \lambda(d', r')\,, \qquad \forall\, (d', r') \in A'\,, \tag{3b}$$

$$w(a')\, x(a') \geq 0\,, \qquad \forall\, a' \in A'\,, \tag{3c}$$

$$t(n') \geq \mathsf{T}(n') - \gamma^-(n')\,, \qquad \forall\, n' \in N'\,, \tag{3d}$$

$$q(n') \leq \mathsf{Q}(n') + \gamma^+(n')\,, \qquad \forall\, n' \in N'\,, \tag{3e}$$

$$x(d', r')\, \big(t(r') - t(d') - \tau_{\mathsf{I}(d'),\mathsf{I}(r')}\big) \geq 0\,, \qquad \forall\, (d', r') \in A'\,, \tag{3f}$$

$$q(r') = t(r') + \tau_{\mathsf{I}(r'),\mathsf{J}(r')}\,, \qquad \forall\, r' \in R'\,, \tag{3g}$$

$$x(d', r')\, \big(q(d') - q(r') - \tau_{\mathsf{J}(r'),\mathsf{J}(d')}\big) = 0\,, \qquad \forall\, (d', r') \in A'\,, \tag{3h}$$

$$\gamma(n') = \gamma^-(n') + \gamma^+(n')\,, \qquad \forall\, n' \in N'\,, \tag{3i}$$

$$\gamma(n') \leq \Gamma\,, \qquad \forall\, n' \in N'\,, \tag{3j}$$

$$\sum_{d' \in D'} x(d', r') \leq 1\,, \qquad \forall\, r' \in R'\,, \tag{3k}$$

$$\sum_{r' \in R'} x(d', r') \leq 1\,, \qquad \forall\, d' \in D'\,, \tag{3l}$$

$$\sum_{n' \in N'} \mathsf{F}(n)\, \gamma(n') + \sum_{a' \in A'} \lambda(a') \leq \mathsf{B}\,, \tag{3m}$$

$$t(n'),\ q(n'),\ \gamma^-(n'),\ \gamma^+(n') \geq 0\,, \qquad \forall\, n' \in N'\,, \tag{3n}$$

$$\lambda(a') \geq 0\,, \qquad \forall\, a' \in A'\,, \tag{3o}$$

$$x(a') \in \{0,1\}\,, \qquad \forall\, a' \in A'\,. \tag{3p}$$

The formulation in model (3) involves nonlinear statements in both the objective function and constraints, which make the ridematching problem intractable to solve, especially for real-size networks. However, there are a few properties of this formulation that help us reduce its size considerably. First, let us emphasize the fact that the objective function in (3a) implies that the unmatched users in any feasible solution will not receive any subsidy. Additionally, this objective function indicates that the optimal solution never assigns any IR incentive to any user, as one unit of IR incentive would decrease the objective function by ϵ. (Note that IR incentives will not be trivially zero when we introduce the ride-back guarantee component in section 4.2). Another observation is that Constraints (3b)-(3i) indicate that the optimal solutions for the incentives and trip start and end times depend on the matching variables. As such, we present a pre-processing procedure in Algorithm 1 that allows us to reduce the problem in (3) to a well-known combinatorial problem.

This algorithm takes the trip information of all users and the original gains of all pairs of drivers and riders as its input. The primary goal of this algorithm is to determine the set of all potential matches in the morning, denoted by A', the optimal values of incentives required for driver d and rider r to have a feasible match, denoted by $\Psi(d', r')$, and the objective coefficient of this pair, denoted by $\mathsf{C}(d', r')$. The algorithm starts with letting A' be an empty set. Next, it iterates over all pairs of riders and drivers to check whether they can be a potential match and if the answer is yes, further find the optimal required subsidies for such pairs. More specifically, in lines 3 to 6, we check the spatio-temporal feasibility of pair (d', r') by solving a linear problem, and storing the optimal trip start times and BA incentives of driver d and rider r conditional on $x(d', r') = 1$, denoted by $\bar{t}(d'|r')$, $\bar{\gamma}(d'|r')$, $\bar{t}(r'|d')$, and $\bar{\gamma}(r'|d')$, respectively. Next, in lines 7, we determine the optimal IR incentive, which is only positive if the corresponding original gain is negative. Finally, we calculate the objective coefficient $\mathsf{C}(d', r')$ as the difference between the original gain and incurred dis-utilities of driver d and rider r from sharing their morning rides together. Also, the total allocated subsidy to (d', r') conditional on them sharing their rides together, denoted by $\Psi(d', r')$, can be computed as the sum of the optimal BA and IR incentives.

Upon this polynomial-time pre-processing procedure, the problem in (3) reduces to a budget-constrained matching problem presented in (5). Note that this formulation clearly implies that those pairs of trips whose IR incentives are positive cannot be a part of the optimal solution.

Algorithm 1: The pre-processing procedure for the morning trips

Input: $\mathsf{I}, \mathsf{J}, \mathsf{T}, \mathsf{Q}, \mathsf{F}, \mathsf{W}$.
Output: A', C, Ψ .

1 Initialize $A' \leftarrow \emptyset$;
2 **for** $(d', r') \in D' \times R'$ **do**
3 Solve the following linear problem:

$$
\begin{aligned}
\min\quad & z = \mathsf{F}(d)\big(\gamma^-(d') + \gamma^+(d')\big) + \mathsf{F}(r)\big(\gamma^-(r') + \gamma^+(r')\big) && && \text{(4a)}\\
\text{s.t.}\quad & t(n') \geq \mathsf{T}(n') - \gamma^-(n') \ , && \forall\ n' \in \{d', r'\}\ , && \text{(4b)}\\
& t(d') + \tau_{\mathsf{I}(d'),\mathsf{I}(r')} \geq \mathsf{T}(r') - \gamma^-(r'), && && \text{(4c)}\\
& t(d') + \tau_{\mathsf{I}(d'),\mathsf{I}(r')} + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} \leq \mathsf{Q}(r') + \gamma^+(r') && && \text{(4d)}\\
& t(d') + \tau_{\mathsf{I}(d'),\mathsf{I}(r')} + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} + \tau_{\mathsf{J}(r'),\mathsf{J}(d')} && && \text{(4e)}\\
& \qquad \leq \mathsf{Q}(d') + \gamma^+(d') && && \\
& \gamma(n') \leq \Gamma \ , && \forall\ n' \in \{d', r'\}\ , && \text{(4f)}\\
& t(d'), t(r'), \gamma^-(d'), \gamma^+(d'), \gamma^-(r'), \gamma^+(r') \geq 0\ . && && \text{(4g)}
\end{aligned}
$$

if (the problem in (4) is FEASIBLE) **then**
4 Retrieve optimal solution $(t^*(d'), t^*(r'), \overset{*}{\gamma^-}(r'), \overset{*}{\gamma^+}(r'), \overset{*}{\gamma^-}(d'), \overset{*}{\gamma^+}(d))$ and objective z^* ;
5 Let $\bar{t}(d'|r')$, $\bar{t}(r'|d') \leftarrow t^*(d')$, $t^*(r')$;
6 Let $\bar{\gamma}(d'|r')$, $\bar{\gamma}(r'|d') \leftarrow \overset{*}{\gamma^-}(d') + \overset{*}{\gamma^+}(d')$, $\overset{*}{\gamma^-}(r') + \overset{*}{\gamma^+}(r')$;
7 Let $\bar{\lambda}(d', r') \leftarrow \max\{0, -\mathsf{W}(d', r')\}$;
8 Let $\mathsf{C}(d', r') = \mathsf{W}(d', r') - (1 + \epsilon)\, z^* - \epsilon\, \bar{\lambda}(d', r')$;
9 Let $\Psi(d', r') \leftarrow \mathsf{F}(d)\, \bar{\gamma}(d'|r') + \mathsf{F}(r)\, \bar{\gamma}(r'|d') + \bar{\lambda}(d', r')$;
10 Update $A' \leftarrow A' \cup \{(d', r')\}$;

$$\max \quad \sum_{a' \in A'} \mathsf{C}(a')\, x(a') \tag{5a}$$

$$\text{s.t.} \quad \sum_{d' \in D'} x(d', r') \le 1, \qquad \forall\, r' \in R', \tag{5b}$$

$$\sum_{r' \in R'} x(d', r') \le 1, \qquad \forall\, d' \in D', \tag{5c}$$

$$\sum_{a' \in A'} \Psi(a')\, x(a') \le \mathsf{B}, \tag{5d}$$

$$x(a') \in \{0,1\}, \qquad \forall\, a' \in A'. \tag{5e}$$

By a reduction from the knapsack problem, the problem in (5) can be shown to be NP-hard. As such, Berger et al. (2011) propose a polynomial time approximation scheme (PTAS) to solve the problem. The core of this PTAS is a Lagrangian relaxation-based method, the solution of which has an objective that is different from the optimal one by at most $2\,\mathsf{C}_{\max}$, where $\mathsf{C}_{\max}$ denotes the largest weight. In Section 4.3, we extend this method to find a near-optimal solution for the problem involving both the morning and evening trips.

4.2 Mathematical Formulation For the Morning and Evening Trips

Let A'' denote the set of all rider-driver trip pairs in the evening, i.e., $A'' = \{(r'', d'') \in R'' \times D''\}$. This set can be generated by applying the pre-processing procedure described in Algorithm 1 to the evening trips. As a result, the ride-matching problem with the morning and evening trips can be formulated as the binary program in (6). The objective function in (6a) seeks to maximize social welfare while ensuring that each dollar spent on subsidy returns more than a dollar in social welfare. Constraints (6b)-(6e) ensure that all users are served at most once both in the morning and in the evening. Constraint (6f) ensures that any rider in R''' is served in the evening if and only if served in the morning. Constraint (6h) sets a limit on the allocated budget.

$$\max \quad \sum_{a \in A' \cup A''} \mathsf{C}(a)\, x(a) \tag{6a}$$

$$\text{s.t.} \quad \sum_{d' \in D'} x(d', r') \le 1, \qquad \forall\, r' \in R', \tag{6b}$$

$$\sum_{r' \in R'} x(d', r') \le 1, \qquad \forall\, d' \in D', \tag{6c}$$

$$\sum_{d'' \in D''} x(r'', d'') \le 1, \qquad \forall\, r'' \in R'', \tag{6d}$$

$$\sum_{r'' \in R''} x(r'', d'') \le 1, \qquad \forall\, d'' \in D'', \tag{6e}$$

$$\sum_{d' \in D'} x(d', r') = \sum_{d'' \in D''} x(r'', d''), \qquad \forall\, r \in R''', \tag{6f}$$

$$\sum_{a \in A' \cup A''} \Psi(a)\, x(a) \le \mathsf{B}, \tag{6g}$$

$$x(a) \in \{0,1\}, \qquad \forall\, a \in A' \cup A''. \tag{6h}$$

Following our discussion in Section 3, we observe that the problem in (6) is similar to that of Lloret-Batlle, Masoud, and Nam (2017) with the addition of Constraint (6g). Thus, by defining a set of auxiliary nodes and edges, we can construct a min-cost max flow network similar to the one explained in Section 3 with the exception that the cost of potential matching edges can be set as $-\mathsf{C}$. Now, if for every potential matching edge we define a new cost, namely the usage fee, and set it to Ψ, problem (6) turns out to be a budget-constrained min-cost flow problem (Holzhauser, Krumke, and Thielen 2016) as follows:

$$\max \quad \sum_{e \in E} \mathsf{C}(e)\, x(e) \tag{7a}$$

$$\text{s.t.} \quad \sum_{e \in \delta^+(v)} x(e) - \sum_{e \in \delta^-(v)} x(e) = 0, \qquad \forall\, v \in N' \cup N'', \tag{7b}$$

$$\sum_{e \in E} \Psi(e)\, x(e) \le \mathsf{B}, \tag{7c}$$

$$x(e) \in \{0,1\}, \qquad \forall\, e \in E \tag{7d}$$

where A''' is the set of auxiliary edges as described in Section 3, $E = A' \cup A'' \cup A'''$, and $\delta^+(v)$ and $\delta^-(v)$ denote the set of in-going and out-going edges of node v in graph G, respectively. It is worth mentioning that, unlike the previous case, the IR incentives are not always equal to zero. Consider the case that $\mathsf{C}(d_1{}', r') < 0$, $\mathsf{C}(r'', d_2{}'') > 0$, and $\mathsf{C}(d_1{}', r') + \mathsf{C}(r'', d_2{}'') > 0$ for arbitrary users d_1, r, d_2. In this case, we may have $x^*(d_1{}', r') = x^*(r'', d''_2) = 1$ which implies that $\lambda^*(d'_1, r') > 0$.

This problem is clearly a generalization of the problem in (5), and thus, can be shown to be NP-hard. Due to the similarities of these two problems, we present a solution methodology that extends the method proposed by Berger et al. (2011) to find near-optimal solutions for large-scale instances of problem (7). This method solves in polynomial time and does not require a commercial optimization engine.

We finalize this section by describing a post-processing procedure to find the optimal trip start time, trip end time, incentives, and gains for the original MINLP based on the optimal matching x^*. For any pair (d', r') such that $x^*(d', r') = 1$, the optimal values of these variables can be calculated as:

$$t^*(r') = \bar{t}(r'|d'), \tag{8a}$$
$$q^*(r') = t^*(r') + \tau_{\mathsf{I}(r'),\mathsf{J}(r')}, \tag{8b}$$
$$\gamma^*(r') = \bar{\gamma}(r'|d'), \tag{8c}$$
$$t^*(d') = \bar{t}(d'|r'), \tag{8d}$$
$$q^*(d') = \bar{t}(d') + \tau_{\mathsf{I}(r'),\mathsf{J}(r')} + \tau_{\mathsf{J}(r'),\mathsf{J}(d')}, \tag{8e}$$
$$\gamma^*(d') = \bar{\gamma}(d'|r'), \tag{8f}$$
$$w^*(d', r') = \mathsf{W}(d', r') + \bar{\lambda}(d', r'), \tag{8g}$$
$$\lambda^*(d', r') = \bar{\lambda}(d', r'), \tag{8h}$$

4.3 A Lagrangian Relaxation-Based Method

Lagrangian Relaxation (LR) is a well-known decomposition technique that proves to be useful in finding an upper bound (in maximization problems) for large MIP problems (Fisher 1981). The MIP problems usually involve some complicating constraints that make them hard to solve. LR takes advantage of such constraints by dualizing them in the objective function and solving an easier subproblem that does not include these constraints. In problem (7), for instance, removing the budget constraint leaves us with a min-cost max flow problem. Since the solution of the LR subproblem is usually infeasible with respect to the relaxed constraint(s), the LR decomposition is usually coupled with a heuristic algorithm to find a high-quality feasible solution. In what follows, we first present the LR subproblem, followed by a patching method that turns the optimal solution of the LR subproblem into a feasible solution with a worst-case optimality bound on the optimal objective of problem (7) in polynomial time. This procedure is summarized in Algorithm 2. For illustrative purposes, consider the small min-cost max flow network G in Figure 2.

4.3.1 The LR Subproblem.

The LR subproblem can be defined by relaxing Constraint (7c) and dualizing it in the objective function with a non-negative Lagrange multiplier, denoted by β, as follows:

$$\max \quad \sum_{e \in E} \mathsf{C}(e)\, x(e) - \beta \,(\sum_{e \in E} \Psi(e)\, x(e) - \mathsf{B}) \tag{9a}$$

$$\text{s.t.} \quad \text{(7b) and } 0 \le x(e) \le 1,\ \forall\, e \in E.$$

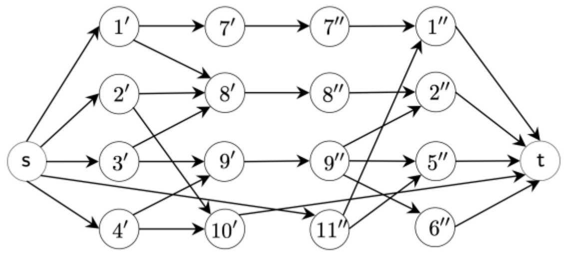


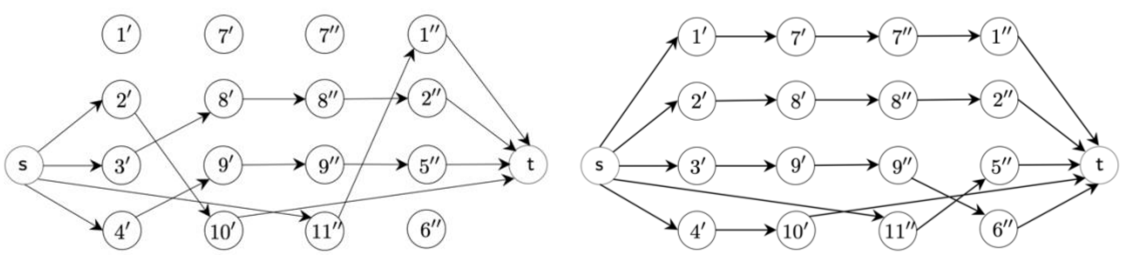
Figure 2: A small example of graph G.

The objective function in (9a) can be simply written as:

$$\mathsf{LR}(\beta) = \sum_{e \in E} \mathsf{C}_\beta\,(e)\, x(e) + \beta\, \mathsf{B}\,, \tag{10}$$

where $\mathsf{C}_\beta(e) = \mathsf{C}(e) - \beta\, \Psi(e)$. Therefore, given the value of β, the cost of the potential matching edge e in graph G can be set to $-\mathsf{C}_\beta(e)$, and the difference between the constant $\beta\,\mathsf{B}$ and the optimal cost of network yields an upper bound for the original problem in (7). We know that LR is a convex piecewise linear function of β, and thus, this upper bound will be minimized at β^*. Also note that the min-cost max flow problem in a directed acyclic graph (DAG) with unit capacities can be solved in strongly polynomial time of $\mathcal{O}(|V|^2|E|)$ using the successive shortest path algorithm (Ahuja, Magnanti, and Orlin 1988). These two facts lead us to the conclusion that β^* can be found in polynomial time of $\mathcal{O}(|V|^4|E|^2)$ using the parametric search method proposed by (Megiddo 1978). Considering the value of $\mathsf{LR}(\beta)$, β^* lies at the intersection of two types of hyperplanes, those with positive slopes induced by budget-feasible flows and those with negative slope induced by budget-infeasible flows. If we let ϵ be a positive infinitesimal number, solving the min-cost max flow in graph G with costs respectively set as $\mathsf{C}_{\beta^*-\epsilon}$ and $\mathsf{C}_{\beta^*+\epsilon}$ gives rise to a budget-infeasible flow denoted by $S_l \subset E$, and a budget-feasible flow denoted by $S_h \subset E$ both of which minimize the LR subproblem. For our small example, for instance, this search can result in two flows S_h and S_l presented in Figures 3(a) and 3(b).

It is easy to show that, unless the two flows are the same, which implies optimality, S_l violates the budget constraint, while S_h satisfies it, and hence, attributes to a feasible solution for the problem in (7). The main issue is that the objective of flow S_h may be arbitrarily far from the optimal one, denoted by OPT. In order to tackle this issue, in the next part, we present a patching method to construct a solution based on two flows S_l and S_h.



(a) S_h (b) S_l

Figure 3: A budget-feasible and -infeasible optimal flow for LR subproblem of the small example.

4.3.2 The Patching Method.

As mentioned earlier, (Berger et al. 2011) proposed a solution method that yields a worst-case optimality of $2\,\mathsf{C}_{\max}$ for a budgeted matching problem. They introduce a patching technique that takes the following steps to derive an approximate solution from Lagrangian optimal solutions S_l and S_h:

1. Turning S_l and S_h into two adjacent extreme solutions such that the difference between the solutions reduces to a single cycle or path.
2. Modifying S_h with a subsequence of the remaining cycle or path using the Gasoline Lemma to yield a budget-feasible high-quality solution.

In this part, we show that this procedure can be generalized to a budgeted min-cost max flow problem. Further, in the next part, we prove that this approach provides us with a worst-case optimality of $3\,\mathsf{C}_{\max}$ for the budgeted min-cost max flow problem. Our generalization is summarized in lines 2-19 of Algorithm 2.

First, let us show how one can turn S_l and S_h into two adjacent extreme flows of the solution polytope of the LR subproblem. Let us define the set of forward edges, E_f, as all edges in S_l that are not in S_h, and the set of backward edges, E_b, as all edges in S_h that are not in S_l. Also, denote the set of reversed edges in E_b by $\overline{E}_b$. Two flows S_l and S_h are adjacent if and only if the graph induced by edges in $\overline{E}_b \cup E_f$, denoted by G', has only one cycle (Gallo and Sodini 1978). The following claim states that $\overline{E}_b \cup E_f$ contains a set of simple paths and cycles.

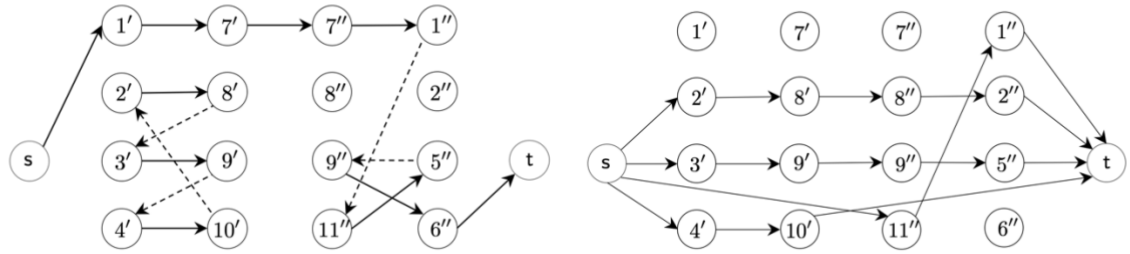
Claim 1: Graph G' contains a non-empty collection of simple path(s) between source s and target t and/or simple cycle(s).

Proof. Based on the definition of $\overline{E}_b \cup E_f$ and the fact that both S_l and S_h satisfy Constraint (7b), we infer that all nodes in V except the source and target node have in-degrees and out-degrees of either 1 or 0. We further infer that the net outflow of s is equal to the net inflow of t. As a result, $\overline{E}_b \cup E_f$ can be decomposed into a combination of cycles and paths that include at least one of the following:

- simple path(s) from s to t ,
- simple path(s) from t to s ,
- simple cycle(s) .

Algorithm 2: A near-optimal solution for P2P ridesharing with ride-back guarantee

Input: $\mathsf{C}, \Psi, G = (V, E)$.
Output: $\tilde{S}$.
1 Find the optimal Lagrangian multiplier β^* and two sets of matching edges S_h and S_l such that $\mathsf{C}_{\beta^*}(S_h) = \mathsf{C}_{\beta^*}(S_l)$ and $\Psi(S_h) \leq \mathsf{B} < \Psi(S_l)$;
2 Let $E_f, E_b \leftarrow S_l \backslash S_h, S_h \backslash S_l$;
3 Find the set of alternating cycles $\mathcal{C}$ in Graph $G' = \left(V, \overline{E}_b \cup E_f\right)$;
4 **repeat**
5 Pick an arbitrary cycle $X \in \mathcal{C}$;
6 Let $X_f, X_b \leftarrow X \cap E_f, \overline{X} \cap E_b$;
7 Let $S \leftarrow \left(S_h \cup X_f\right) \backslash X_b$;
8 **if** $\Psi(S) \leq \mathsf{B}$ **then**
9 Update $S_h \leftarrow S$;
10 **else**
11 Update $S_l \leftarrow S$;
12 Update $E_f, E_b \leftarrow S_l \backslash S_h, S_h \backslash S_l$;
13 Find the set of alternating cycles $\mathcal{C}$ in Graph $G' = \left(V, \overline{E}_b \cup E_f\right)$;
14 **until**$(|\mathcal{C}| = 1)$
15 **if** $\Psi(S_h) = \mathsf{B}$ **then**
16 Let $\tilde{S} \leftarrow S_h$;
17 **else**
18 Find sequence $Y = Y_f \cup Y_b$ in cycle X using the Gasoline Lemma ;
19 Find $\tilde{S}$ by solving a min-cost max flow in graph $G'' = \left(V, \left(S_h \cup Y_f\right) \backslash Y_b\right)$ with costs C ;



(a) $\overline{E}_b \cup E_f$ (b) S_h after cancelling a cycle

Figure 4: The cycles in $\mathcal{C} = \{X_1, X_2\}$ and the flow S_h after cancelling cycle induced by the edges in $X_1 = \{(2', 8'), (8', 3'), (3', 9'), (9', 4'), (4', 10'), (10', 2')\}$. In part (a), the solid arrows denote edges in E_f and the dashed arrows denote the edges in $\overline{E}_b$.

Note that a path from s (t) to t (s) in graph G' can turn into a cycle using an auxiliary edge from t (s) to s (t) with cost 0 and capacity 1. Thus, we conclude that graph G', with minor adjustments, consists of a non-empty set of cycles denoted by $\mathcal{C}$. If the cardinality of $\mathcal{C}$ is one, by definition S_l and S_h are adjacent. Otherwise, until the cardinality of $\mathcal{C}$ is equal to one, we repeat the process described in lines 4 to 13 of Algorithm 2. More specifically, we draw one arbitrary cycle from $\mathcal{C}$ at a time and add its forward edges to S_h and remove its backward edges from S_h. This yields a new feasible flow, denoted by S, the Lagrangian objective of which is equal to those of S_h and S_l. Now, depending on whether this flow satisfies the budget constraint, we replace it by S_h or S_l. At the end of this process, if flow S_h depletes all the budget, we claim that it is the optimal solution. Otherwise, we continue to the patching method described below. In Figure , we show that $\mathcal{C}$ consists of two cycles $X_1 = \{(2', 8'), (8', 3'), (3', 9'), (9', 4'), (4', 10'), (10', 2')\}$ and $X_2 = \{(\mathsf{s}, 1'), (1', 7'), (7', 7''), (7'', 1''), (1'', 11''), (11'', 5''), (5'', 9''), (9'', 6''), (6'', \mathsf{t})\}$ in our small example. Note that X_2 is indeed a path from s to t which can turn into a cycle by adding the auxiliary edge (t, s). Since $|\mathcal{C}| = 2$, we need to cancel one of these cycles using the procedure described above. Note that the forward edges in X_1 and X_2 are shown by solid lines and backward edges are shown by dashed lines. In Figure , we show the result of cancelling cycle X_1 which provides a new flow S_h, assuming that the budget constraint remains feasible.

Now, let us denote the last remaining cycle in $\mathcal{C}$ by $X = X_f \cup \overline{X}_b$, where X_f and $\overline{X}_b$ respectively denote the forward edges and backward edges in X. If we use the procedure described above to remove the last cycle X, flow S_h will turn into flow S_l, i.e., the two solutions will collapse, providing an infeasible solution with respect to the budget constraint. Also note that:

$$\mathsf{C}(S_h) \geq \mathsf{OPT} - \beta^* \left(\mathsf{B} - \Psi(S_h)\right).$$

Therefore, finding a solution whose subsidy is closer to B improves the optimality bound of our solution for problem (7). To this end, we present a method to alter flow S_h using a subsequence of edges in cycle X. Similar to the patching method in (Berger et al. 2011), the core of our method is based on the Gasoline Lemma, presented in Appendix B.

In order to apply the Gasoline Lemma, we define the following weights for edges $e = (i, j)$ in X:

$$\alpha(i, j) = \begin{cases} \mathsf{C}_{\beta^*}(i, j), & \text{if } (i, j) \in X_f, \\ -\mathsf{C}_{\beta^*}(j, i), & \text{if } (i, j) \in X_b. \end{cases} \tag{11}$$

For any rider $r \in R'''$ in X, we further adjust the weight of the edge that starts from or ends at r'' as follows:

$$\alpha(r'', j) = \alpha(r'', j) + \max\{\mathsf{C}(d'_l(r'), r'),\ \mathsf{C}(d'_h(r'), r'),\ 0\}, \tag{12}$$

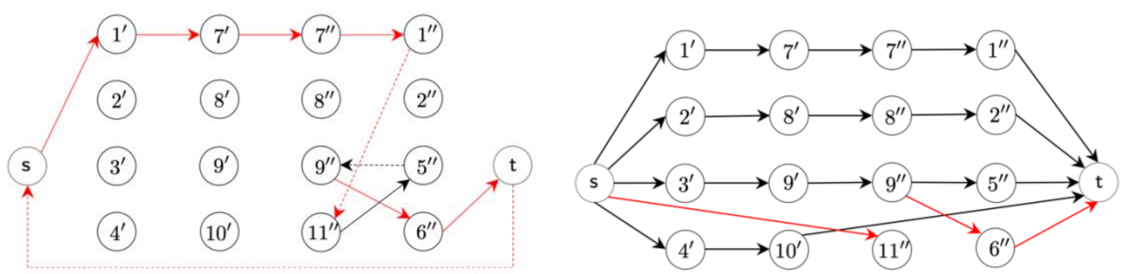
$$\alpha(i, r'') = \alpha(i, r'') - \max\{\mathsf{C}(d'_l(r'), r'),\ \mathsf{C}(d'_h(r'), r'),\ 0\}, \tag{13}$$

where $d'_l(r')$, $d'_h(r')$ respectively denote the driver trips that are matched with rider trip r' in flows S_l and S_h. If there is no such matches in either of these flows, we let the corresponding C value be zero. Also, note that the α values for edges in X add up to zero. Next, we apply the Gasoline Lemma to the edges in X with the weights set as α. This lemma gives us an edge e_1 in X such that any subsequence X' of X that starts with edge e_1 has the following property:

$$\sum_{e \in X'} \alpha(e) \geq 0, \tag{14}$$

Let $Y = Y_f \cup \overline{Y}_b$ be the longest subsequence X' such that adding the edges in Y_f to S_h and removing the edges in Y_b from it do not violate the budget constraint. We further remove edges from the start and end of Y that originate from or end at the source and target nodes, respectively. (Note that doing so does not affect the property in (14) since the value of α is zero for all such edges.) Let S denote the flow built by adding edges in Y_f to S_h and removing the edges in Y_b from it. Note that flow S satisfies the budget constraint due to the choice of Y. However, it may not attribute to a feasible flow as it may violate the flow conservation constraints in (7b). Therefore, we solve a min-cost max flow in graph $G'' = (V, S)$, with costs of edges set as C, to obtain a feasible flow $\tilde{S}$. This implies that $\tilde{S} \subseteq S$.

Let us implement this procedure for our small example. In Figure 5(a), we show the remaining cycle $X = X_2$ after removing cycle X_1. We further show a possible case for subsequence Y of X, which is represented by the edges in red, starting from node $9''$ and ending at node $11''$. Figure 5(b) demonstrates graph $G' = (V, S)$. Note that S does not satisfy the flow conservation for nodes $9''$ and $11''$. By solving a min-cost max flow in graph G, we obtain flow $\tilde{S}$ which can be obtained by removing the edges in red from flow S.



(a) X and Y (b) $\tilde{S}$

Figure 5: Cycle X, subsequence Y, graph $G' = (V, S)$ and the final flow $\tilde{S}$

4.3.3 Properties of the LR-based Method.

In this part, we show that $\tilde{S}$ yields a worst-case optimality bound, and solves in strongly polynomial time.

Proposition 1: *Algorithm 2 provides a solution for problem (7) with worst-case optimality bound of* $3\,C_{max}$, *i.e.* $C(\tilde{S}) \geq OPT - 3\,C_{max}$.

Remark 1: *Algorithm 2 provides a solution for the problem in (7) in strongly polynomial time.*

Proof. As mentioned earlier, the value of β^* can be found in $\mathcal{O}(|V|^4|E|^2)$ using the Megiddo's parametric search technique. Also, all simple cycles in graph G' can be found in $\mathcal{O}(|V|^2)$ using the algorithm proposed by Johnson (1975), given the fact that at most $\mathcal{O}(|V|)$ simple cycles can exist in this graph. After each loop, at least one cycle can be cancelled. Thus, the whole cycle-cancelling process takes $\mathcal{O}(|V|^3)$. The Gasoline Lemma can be applied in $\mathcal{O}(|V|)$, and flow $\tilde{S}$ can be found by solving a min-cost max flow in $\mathcal{O}(|V|^3)$, given the fact $\mathcal{O}(|E|) = \mathcal{O}(|V|)$ in graph G''. Thus, the running time of Algorithm 2 is bounded by that of the parametric search and the result follows.

We finalize this section by stating that one can turn Algorithm 2 into a PTAS by guessing the heaviest edges that will be in the optimal solution similar to the procedure described by Berger et al. (2011). However, in large-scale instances of our problem $\mathsf{C}_{\max}$ is sufficiently small compared to OPT, which implies that the optimality gap is reasonably low and we do not need such an expensive procedure.

4.4 A Budget-Balanced Variant of the Incentive Program

In Section 3, we assumed that the ridesharing system relies on an external budget of size B to incentivize the participants. In this subsection, we introduce a budget-balanced variant of our traveler incentive program whose funding comes from taxing the matches with positive adjusted gain. As such, let ϕ denote a flat tax rate that will be applied to any matched pair whose gain after adding subsidies is positive. In this case, we can replace the budget constraint in (7c) with the following constraint:

$$\sum_{e\in E} \Psi(e)\, x(e) \ \leq \phi \sum_{e\in E} (\mathsf{W}(e) + \bar{\lambda}(e))\, x(e)\,. \tag{15}$$

In problem (7), we chose to maximize the difference between system's social welfare and allocated subsidy to ensure that every external dollar added to the system generates a positive return on investment. In the case of a budget-balanced system, however, the budget will be provided internally, and thus, we choose to maximize the system's social welfare. As a result, the after-tax social welfare of the budget-balanced system can be calculated:

$$(1-\phi) \sum_{e\in E} (\mathsf{W}(e) + \bar{\lambda}(e))\, x(e)\,, \tag{16}$$

Note that the pre-processing procedure in Algorithm 1 is still valid in this case, since taxing affects neither the spatio-temporal feasibility nor the individual rationality of any pair. Hence, given a constant value for ϕ, we can find the optimal solution to the new problem using the methodology proposed in the previous subsection with minor adjustments.

For a budget-balanced system, an important question that needs to be addressed is that "what would be an appropriate value for ϕ?". On the one hand, increasing the value of ϕ raises our available budget to subsidize more users, which leads to a higher matching rate and system-level social welfare. On the other hand, raising the tax rate adversely affects the social welfare of those participants who could be matched with lower values of ϕ. In what follows, we present a proposition that helps us answer this question.

Proposition 2: *The optimal flat tax rate that yields the maximum social welfare for the budget-balanced variant of the traveler incentive program can be found as:*

$$\phi_{max} = \frac{\sum_{e\in E} \Psi(e)\, x_{max}(e)}{\sum_{e\in E} (W(e) + \bar{\lambda}(e))\, x_{max}(e)}\,, \tag{17}$$

where x_{max} denotes the optimal solution to problem (7) when Constraint (7c) is relaxed.

Proof. Let us rewrite the objective function in (16) as:

$$\max \qquad \sum_{e\in E} \mathsf{C}(e)\, x(e) + (\sum_{e\in E} \Psi(e)\, x(e) - \phi \sum_{e\in E} (\mathsf{W}(e) + \bar{\lambda}(e))\, x(e))\ .$$

Note that the statement inside the parentheses is always non-positive due to the budget constraint in (15). Therefore, its value will be maximized if Constraint (15) is binding. Thus, solution $(x_{\max}, \phi_{\max})$ maximizes both the statements inside and outside the parentheses and the result follows. This proposition further implies that the budget-balanced variant of our program is no longer NP-hard as stated in the following remark.

Remark 2: *The optimal solution to the budget-balanced variant of the traveler incentive program as a function of ϕ can be found in strongly polynomial time.*

Proof. Since problem (7) without Constraint (7c) can be modeled as a min-cost max flow and thus the worst-case running time complexity of finding $x_{\max}$ is $\mathcal{O}(|V|^2|E|)$. Also, the pre-processing procedure has a time complexity of $\mathcal{O}(|V|^2)$. Hence, the overall complexity of the budget-balanced variant of the incentive program is the same as that of the min-cost max flow problem which is fully polynomial in the input size and the result follows.

5 Numerical Experiment

In this section we conduct a comprehensive set of numerical studies using the New York City taxi dataset (http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml) to evaluate the performance of the proposed traveler incentive program for a ridesharing system. In the next three subsections, we carefully define our dataset, parameter settings, and several performance metrics that help us quantify the impact of the proposed methodology. Afterwards, we present the result of various experiments.

All the experiments are implemented on a 3.50 GHz Intel Xeon machine with a 64-bit version of the Windows 10 operating system with 128.0 GB RAM. The data preparation and the LR-based algorithm are coded in `Python 3.7`, and all the optimization problems are solved using `GUROBI 9.0`.

5.1 Dataset

In our numerical experiments, we assume that the proposed ridesharing system is practiced by the travelers in the Manhattan area of the New York City. Therefore, the road network of the Manhattan area is extracted from `Open Street Map`, which consists of 4500 nodes (stations) and 9800 transportation links. We further use the `Google Map API` to find the shortest-path travel time and driving distance between every pair of stations in the morning and evening. For the morning and evening trip information of the participants, we use the daily average of the historical trips in the New York City taxi dataset from Feb 1, 2016 to Feb 10, 2016 subject to the following assumptions and modifications:

- For a user that participates in the system only in the morning or evening peak hours, the origin and destination stations as well as the desired earliest departure time are obtained directly from the dataset.
- To construct the set of users who are interested in both the morning and evening trips, we pick the trip information of their morning trips from the dataset. For the evening trip, we randomly pick a desired earliest departure time from the evening peak hour period and flip the origin and destination of their morning trip to form the origin and destination of the evening trip.
- In order to obtain a tight time window for every trip, we generate a random number following a Normal distribution with mean of 5 minutes and standard deviation of 1 as the trip time flexibility. As a result, the desired latest arrival time of every trip can be set as the sum of desired earliest departure time, the shortest-path travel time between their origin and destination stations, and the trip time flexibility.
- We only consider trips whose shortest-path travel times are at least 10 minutes.
- We randomly generate the user's values of time and distance from Normal distributions with means $\mu(\mathsf{F})$ and $\mu(\mathsf{H})$ and standard deviations $\mu(\mathsf{F})/5$ and $\mu(\mathsf{H})/5$, respectively.
- For every trip, the maximum allowed extension in the time window, Γ, is set to 15 minutes.

5.2 Experiment setup and Parameter Settings

In our numerical experiments, we consider different scenarios of the ridesharing network with parameters defined as follows:

- $|N|$: the total number of participants in the system,
- $|R|/N$: the percentage of riders,
- $|N'''|/N$: the percentage of participants with both morning and evening trips,
- B : the total available budget for subsidy,
- $\mu(\mathsf{F})$: the mean of a user's time valuation,
- $\mu(\mathsf{H})$: the mean of a user's distance valuation,
- M and E : the morning and evening peak hour periods, respectively.

We define a base scenario whose parameters are shown in Table 1. We define other scenarios by changing the value of a single parameter in the base scenario at a time.

Table 1: The parameter settings for the base scenario.

$\lvert N \rvert$	$\lvert R \rvert / N$ (%)	$\lvert N''' \rvert / N$ (%)	B ($)	μ(F) ($/min.)	μ(H) ($/mile)	M (am/am)	E (pm/pm)
6000	50	50	1000	0.35	3.0	$[7{:}00, 10{:}00]$	$[5{:}00, 8{:}00]$

Let us first describe how we generate the different sets, used in developing our methodology, based on the value of the first three parameters in Table 1. Consider the base scenario where the set of all users can be shown as $N = \{1, \ldots, 6000\}$ based on the first parameter. The set of all riders and drivers can be respectively shown as $R = \{1, \ldots, 3000\}$ and $D = \{3001, \ldots, 6000\}$ based on the value of 50% for the second parameter. Finally, the value of 50% for the third parameter allows us to find all other sets as $R' = \{1', \ldots, 2250'\}$, $R'' = \{751'', \ldots, 3000''\}$, $R''' = \{751, \ldots, 2250\}$, $D' = \{3001', \ldots, 5250'\}$, $D'' = \{3751'', \ldots, 6000''\}$, and $D''' = \{3751, \ldots, 5250\}$. We also set ϵ in the parametric search to 10^{-6}.

5.3 Performance Metrics

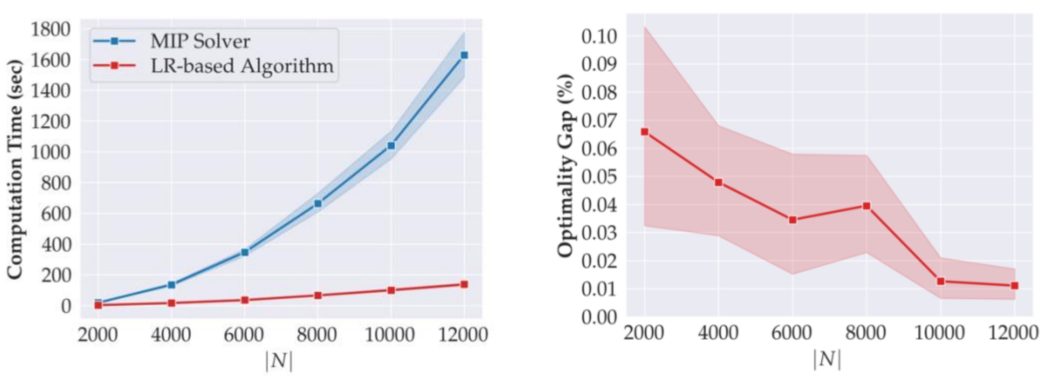
In order to quantify the performance of the proposed methodology, we introduce the following performance metrics:

- Social Welfare (SW) = $\sum_{e \in E} w^*(e)\, x^*(e)$,
- Subsidy Impact Rate (SIR) = (SW $-$ SW without Subsidy)$/(\sum_{e \in E} \Psi(e)\, x^*(e))$,
- Subsidized Matches Rate (SMR) = 100 × $(\sum_{e \in E} 1_{\Psi(e)>0})/(\sum_{e \in E} x^*(e))$,
- Time Window Extension (TWE) = $(\sum_{n \in N' \cup N''} \gamma^*(n))/(\sum_{n \in N' \cup N''} 1_{\gamma^*(n)>0})$,

where "Social Welfare without Subsidy" and "Matching Rate without Subsidy" respectively represent the negated total flow cost and matching rate of the original min-cost flow problem in Section 3 without considering any subsidy.

5.4 The Performance of the LR-Based Solution Method

In this study, we show the merits of our proposed solution method in solving the budget-constrained min cost flow problem in (7). To this end, we compare the computation time and the quality of the solution obtained by the proposed LR-based method with that of solving the problem in (7) directly using a MIP solver. For this comparison, we generated 10 random instances of the problem using the parameters in the base scenario and applied both methods to solve them. We also repeated our experiments with the number of users set as 2000, 4000, 6000, 8000, 10000, and 12000. Figure shows the computation time of both methods for different number of users averaged over 10 instances. Figure shows the average optimality gap of our solution method for different number of users. The shaded region in both figures demonstrate the 95% confidence intervals of the average values.



(a) Computation time (b) Optimality gap

Figure 6: The comparison between the performance of the MIP solver and the LR-based Method

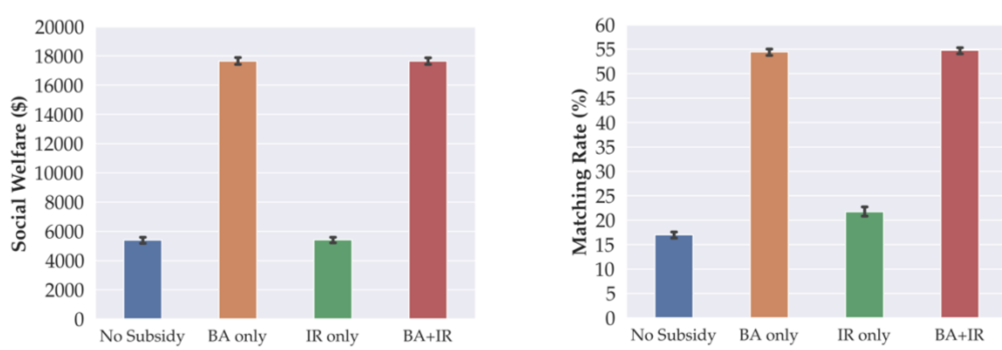
Figure (6a) indicates that the MIP solver takes almost half an hour to solve the problem with 12000 customers while the proposed method in Algorithm 2 takes less than 3 minutes to find a near-optimal solution. Also, this figure clearly shows that the rate of increase in the computation time of the MIP solver is super-linear while the computation time of our solution method increases fairly linearly with the number of users, highlighting the scalability of the proposed method. Also, Figure 6(b) suggests that, in general, the quality of the solutions obtained using Algorithm 2 improves with the number of system participants. Finally, it is worth mentioning that we observed that the MIP solver uses different heuristics to solve the problem in the reported times. Without these heuristics, the pure Branch-and-Cut algorithm takes a considerable amount of memory and hours of computation time to provide any high quality solution to any instance of the problem with more than 6000 users.

5.5 The Impact of the Two Proposed Incentives

In this paper, we consider two types of incentives, namely the behavioral adjustment and the individual rationality incentives. Here, we compare the performance of a ridesharing system with guaranteed-ride-back under four different cases of "no subsidy", "only the BA incentive", "only the IR incentive", and "both the BA and IR incentives" using 10 different randomly-generated instances of the base scenario. Figure demonstrates the average system-level social welfare over 10 runs for these four cases. Clearly, no subsidy yields the lowest values in social welfare. Subsidizing the system only using the BA incentive type shows a promising increase in the social welfare. However, this figure suggests that the IR incentive does not change the social welfare in the absence of the BA incentive, and improves the social welfare in the presence of the BA incentive only minusculely. There are two main reasons for this observation. First, the IR incentive cannot turn a spatio-temporally infeasible pair into a feasible pair by itself. That is the reason why in Algorithm 1 we first check the spatio-temporal feasibility of a pair, and only when a pair is spatio-temporally feasible do we inspect the individual rationality of that pair. Secondly, we stated in Section 4.2 that the IR incentive can be positive only if the served rider is in R''' and one of their trips in the morning or evening has negative original gain and the other has positive original gain, with sum of their original gains being positive. Otherwise, the IR incentive cannot add a higher value than its magnitude to the system. This fact highly narrows down the number of edges with positive IR incentive, and thus, its impact on the system-level social welfare diminishes.

From the experiments above, one may come to the conclusion that the IR incentive seems to be useless and should not be considered. In order to show the merit of such an incentive in promoting a ridesharing system, we change the objective function in problem (7) to maximizing the matching rate as follows:

$$\max \quad \frac{100 \times 2}{|N'| + |N''|} \sum_{e \in E} x(e) \, .$$



(a) Maximize the social welfare (b) Maximize the matching rate

Figure 7: The comparison between the performance of the ridesharing under different cases of subsidy for the objective of (a) maximizing the social welfare, and (b) maximizing the matching rate.

It is easy to show that under this objective, we can still use the pre-processing procedure in Algorithm 1. Figure demonstrates the result of maximizing the matching rate under the four cases. This figure indicates that the IR incentive can increase the new objective, i.e., the matching rate, in a statistically significant manner compared to the no subsidy case, even in the absence of the BA incentive. However, this figure also suggests that the BA incentive has more benefits than the IR incentive to offer due to the first reason discussed above. This figure also shows that in the presence of the BA incentive, introducing the IR incentive does not add considerable value.

5.6 The Impact of Tax Rate in the Budget-Balanced Incentive Program

In Section 4.4, we introduced a budget-balanced variant of our incentive program in which the required subsidy is collected internally through taxing the matches with positive adjusted gain. We further showed that the optimal flat tax rate $\phi_{\max}$ can be found analytically. In this experiment, we demonstrate the empirical results for an arbitrary instance of the base scenario for different values of $0 \leq \phi \leq 100$. Figure 3 shows the after-tax social welfare in (16) against the tax rate. This figure clearly shows that the social welfare of the budget-balanced variant of the incentive program increases sub-linearly until it reaches $\phi_{\max} = 16.87\%$, and then it starts to collapse until it gets to zero at 100%. Overall, this figure suggests that subsidizing the system internally can significantly increase its social welfare from less than \$6000 (without subsidy) to more than \$20,000. Also, it indicates that even a small tax rate as low as 1% can double the system's social welfare.

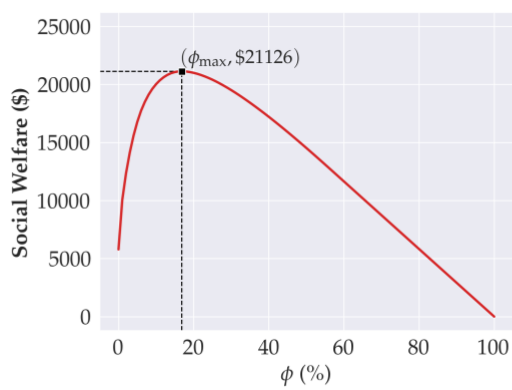
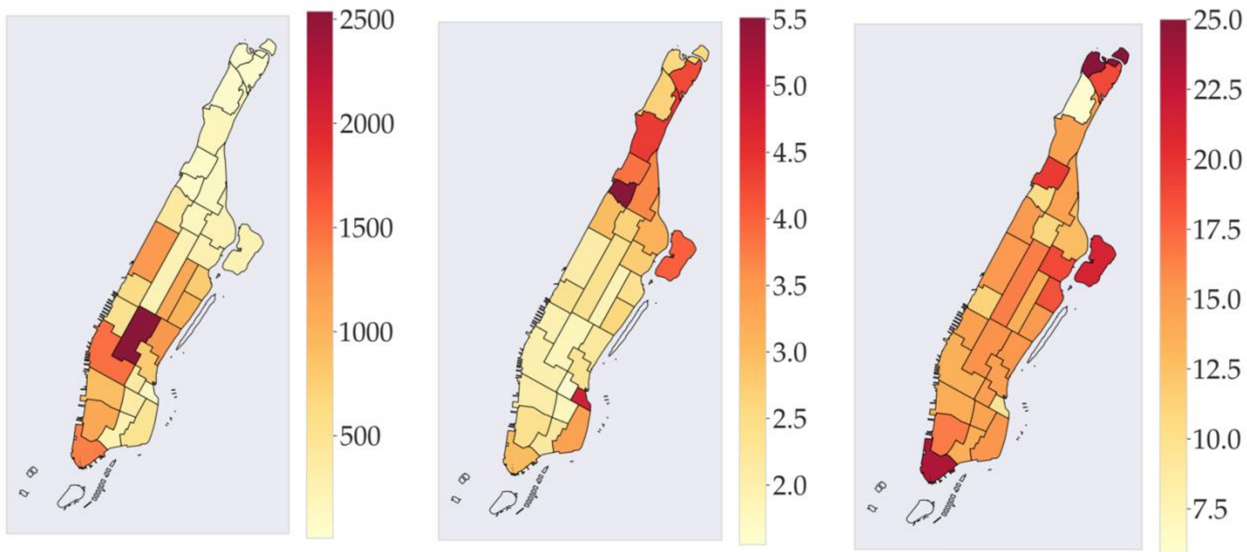


Figure 8: The social welfare for different values of tax rate in a budget-balanced incentive program

5.7 The Distribution of the BA Incentive Based on Trip Origins and Destinations

In this experiment, we are interested in learning the characteristics of the users who are more likely to change their travel behavior, and thus, receive the BA incentive. To this end, we investigate the chance of receiving the BA incentive as well as the average amount of subsidy based on the origin and destination geo-coordinates of participants. Figure 9 presents three heat maps that help us analyze the distribution of the BA incentive among participants for an arbitrary instance of the base scenario. Note that due to having a large number of stations, we aggregate them into larger zones that represent different neighborhoods in the Manhattan area.

Figure 9(a) shows the number of trips in the morning and evening whose origin or destination stations falls within each zone. This figure clearly shows that most trips originate/end form/at the lower Manhattan area, more specifically in the neighborhood where the Time Square is located. Figure (9b) shows the percentage of the subsidized users for each neighborhood. This figure suggests that those users who start or end their trips in less popular zones are more likely to receive the BA incentive. This is not surprising, because the chances of finding users with compatible trips is lower for such trips. Therefore, they are more likely to extend their time windows to match with other users. Figure (9c) shows the average amount of the BA incentive per number of users that received such an incentive for each neighborhood. This figure clearly indicates that as the popularity of a neighborhood decreases, the trips whose origin and/or destination stations fall into that region should expand their time windows by a larger extent. Overall, this experiment suggests that those users who are located in the areas that are poorly supported by public transit can benefit more from the introduction of such an incentive program.



(a) Frequency of the Origin and Destination Zones

(b) Relative frequency of subsidized zones

c) Average subsidy per subsidized users of zones

Figure 9: The heatmap of the origin and destination zones

5.8 A Dynamic Implementation of the Incentive Program

Throughout this paper, we considered a static ridesharing system for which all trip information for a given day is known prior to solving the ride-matching problem. This requires all users to announce their trips before the start of the morning period. In this experiment, we relax this requirement by considering a dynamic system in which users are allowed to announce their trip information shortly before their desired earliest departure times. Let $\mathsf{L}(n)$ denotes the trip announcement time for user $n \in N$. Note that we assume that users in N''' announce both of their trips at the same time in the morning. In order to solve the dynamic ride-matching problem, we adopt the rolling horizon approach that re-optimizes the static ride-matching problem in short time intervals, say every 1 minute, using all the available trip information collected by the onset of each interval (see e.g., Agatz et al. (2011)). There are two important considerations for applying the rolling horizon approach in the existence of an incentive program: (i) how to allocate the total budget B to the re-optimization intervals, and (ii) when to finalize the matches between riders and drivers. Let us denote the re-optimization points (i.e., the starting times of the re-optimization intervals) by set $P = P' \cup P''$, where P' and P'' respectively denote the re-optimization points in the morning and evening. It is easy to show that the pre-processing (with a minor adjustment) and solution method can still be applied in this case. More specifically, we only need to add the following constraint to the problem in (4):

$$t(n') \geq p + 1, \qquad \forall\, n \in d, r\,.$$

This constraint ensures that the trip's start time is always after the end of the re-optimization interval. In what follows, we introduce two different policies for each of these considerations and evaluate their impacts on the system's social welfare.

For allocating the budget to the intervals, we consider two policies: (1) a naive policy that allows for using the entire available budget in each interval, and (2) dividing the total budget among intervals in a two-step process. In the first step, we divide the total budget between the morning and the evening intervals according to a pre-determined rate $\pi \in (0,1)$. In the second step, we divide $\pi\,\mathsf{B}$ dollars uniformly among the intervals in P', and $(1-\pi)\,\mathsf{B}$ dollars uniformly among the intervals in P''. Also, we roll over the unused budget in each interval to the next interval.

For finalizing the matches found after solving the problem in [eq:P] at each time $p \in P$, we consider two policies: (1) we finalize the matches for users as soon as they got matched in a re-optimization interval, (2) we postpone the matching finalization for any user as much as possible. More precisely, in the second policy we only finalize the matching between pairs of riders and drivers who cannot get matched together in the next re-optimization interval, or when the amount of subsidy allocated to them needs to increase in the next interval. This can be done by solving the problem in (4) once for $p + 1$, and once for $p + 2$ at every re-optimization point $p \in P$.

Figure 4 demonstrates the results of applying these policies to 10 randomly-generated instances of the base scenario. For each scenario, the trip announcement times are calculated by subtracting a random number following the Normal distribution with mean 5 and standard deviation of 1 from the corresponding desired earliest departure times. Moreover, for partial budget allocation policy, we consider different values of π from 0.5 to 1. The scenario "whole" in this figure corresponds to the naive budget allocation strategy of using as much of the budget as desired in the earlier intervals, and only using the roll-over budget in the subsequent ones. Scenarios "Partial$_\pi$" indicate the results for the fraction π of the total budget being allocated to the morning period. Under each scenario, the blue and red bars indicate the total social welfare for the two match fixing policies of finalizing the matches as early and late as possible, respectively.

Figure 10 clearly indicates that the combination of a partial budget allocation and latest time to finalize the matches yields the highest system-level social welfare. This is due to the following facts: under the Whole budget policy, the available budget will be depleted in the first few re-optimization points, which results in a huge opportunity cost of not being able to subsidize the ride-matching problems in the following intervals. Additionally, postponing the matching finalization to the latest possible interval enables us to obtain more information on future trips, and thus, find possibly more beneficial matches (with higher welfare and lower subsidy) for each user. This figure further demonstrates that the highest social welfare can be obtained by setting π to 0.8, which is slightly higher than the rate of the morning trips in the system (0.75 in the base scenario) in the base scenario. One possible explanation for this observation is that all the available budget will not be used in each interval, and hence, letting π be slightly larger than the rate of the morning trips in the system can help us distribute the total available budget more uniformly between the morning and evening trips; i.e., if there is no need for the extra 5% of budget, it will be rolled over to the next interval, but not proving the possibly of using this budget may lead to an opportunity cost.

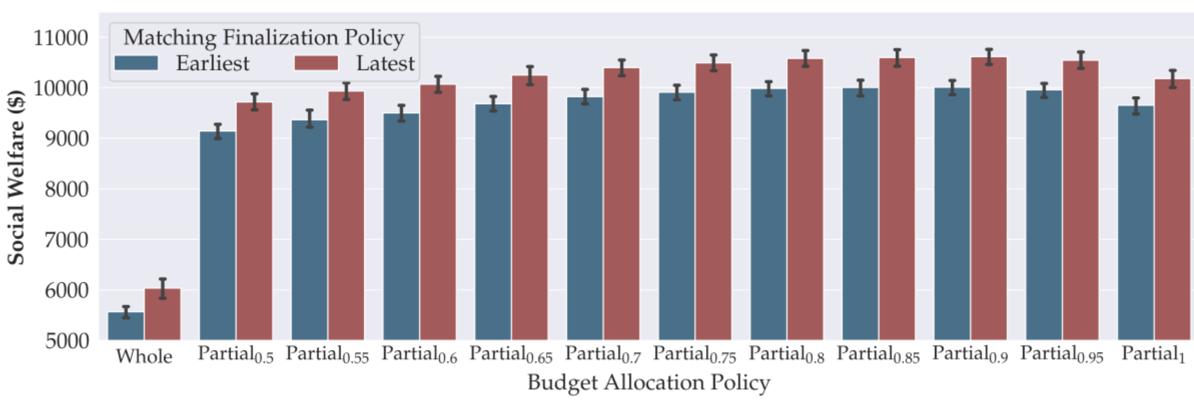
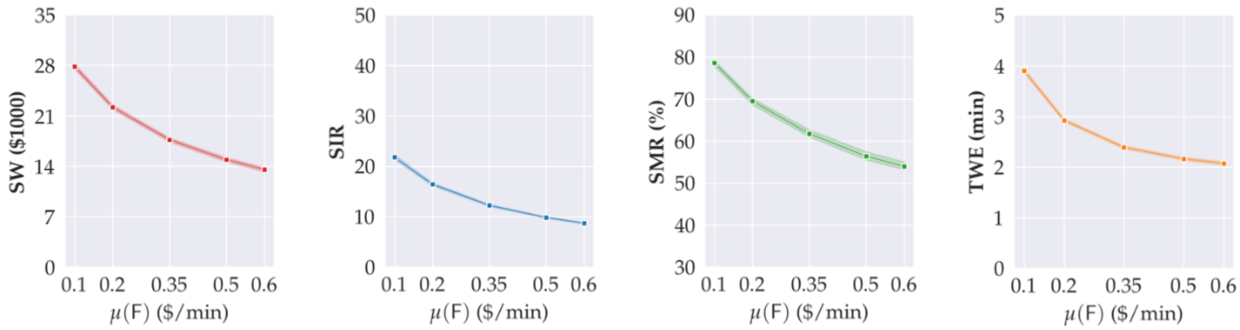


Figure 10: The social welfare a dynamic incentive program under different policies for allocating bidget and finalizing matches.

5.9 Sensitivity Analysis

In this subsection, we consider different scenarios, by changing one parameter of the base scenario at a time, to analyze the sensitivity of the performance metrics (see Section 5.3) on parameter values (see Section 5.2). The results of this analysis are plotted in Figures 11, 12, D.1, D.2, D.3, D.4, and D.5. Each plot shows the average and the 95% CI of the corresponding performance metric over 10 instances. In what follows, we analyze the results of Figures 11 and 12. The rest of Figures and their interpretations are presented in Appendix D.

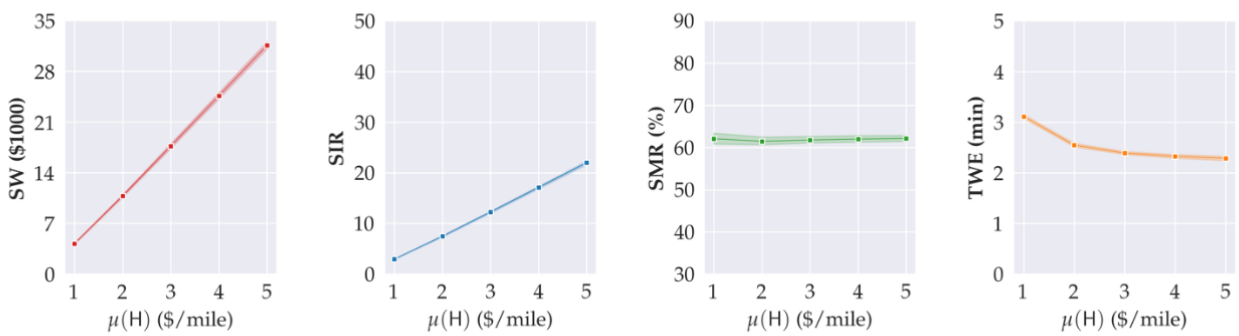
In the base scenario, we assumed that the user's value of time, $\mu(\mathsf{F})$, has a mean of \$0.35 per minute. In order to study the relation between different values of this parameter and the performance metrics, we change the value to \$0.1, \$0.2, \$0.5 and \$0.6 per minute. The results are presented in Figure 11. In this figure, we observe that all the performance metrics decrease sublinearly with an increase in the value of $\mu(\mathsf{F})$. This is not surprising because the amount of subsidy allocated to a user to change their travel behavior is a linear function of their value of time. Therefore, as a user's value of time increases, the system has to compensate them with more subsidy to change their travel behavior, i.e. expand their time window.



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure 11: Impact of average value of time

Next, we consider the impact of changes in the average user's value of distance, denoted by $\mu(\mathsf{H})$, on the performance metrics. Figure 12 presents the results of changing this value from \$3 per mile in the base scenario to \$1, \$2, \$4 and \$5 per mile. Figure 12(c) shows no significant changes in the percentage of subsidized matches, which is due to the fact that the BA incentive allocated to users does not get affected by the value of this parameter, and that most of the available budget is spent on the BA incentive (see Section 5.5). However, Figures 12(a) and 12(b) indicate that both the social welfare and the subsidy impact rate increase linearly as a result of an increase in the value of $\mu(\mathsf{H})$. The main reason behind these increasing trends is that an increase in the valuation of distance linearly increases the savings due to matching, which leads to a higher number of individually rational matches. Also, this increase in the number of individually rational matches allows users to find matches that not only have higher gains but also require smaller changes in their time windows, as shown in Figure 12(d).



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure 12: Impact of average value of distance

6 Conclusion and Future Work

In this paper, we consider a community-based ridesharing system with guaranteed-ride-back for commuters. To promote such a ridesharing system among commuters, we introduce a traveler incentive program (TIP) that offers two types of incentives, namely the behavioral adjustment (BA) and the individual rationality (IR) incentives. We formulate the joint problem of matching, scheduling, and incentive allocation as a mixed integer nonlinear program (MINLP). Using a pre-processing procedure and by utilizing linear programming, we reduce the MINLP problem to a budget-constrained min-cost flow problem. For solving large-scale instances of this problem, we devise a polynomial-time Lagrangian Relaxation-based algorithm, and obtain a worst-case optimality bound for its performance. We further introduce a budget-balanced variant of the incentive program that does not require external budget. Finally, we conduct several experiments using the New York City taxi dataset to evaluate different aspects of the TIP and the solution methodology. Our findings from the results of these experiments can be summarized as follows:

3. The proposed TIP considerably increases the social welfare by tripling that of the system without subsidy. Also, every dollar spent on subsidy increases the social welfare by 12 dollars and matching rate by more than 40%.
4. The proposed LR-based solution method significantly reduces the computational effort needed to solve the large-scale instances of the problem when compared to a MIP solver. Also, the relative optimality gap of the LR method does not exceed 0.15% for any larges-scale instance, and converges to zero as the size of the problem increases.
5. When the objective of the system is to maximize social welfare, the impact of the IR incentive is negligible. However, we show that this type of incentive can be effective when the system operator aims to maximize the matching rate. Nonetheless, when BA incentives are introduced, adding IR incentives to the incentive basket improves neither the social welfare nor the matching rate in a statistically significant manner. This signifies the importance of using limited resources to change travel behavior, rather than providing direct monetary subsidies to travelers.
6. Those users who start or end their trips in regions that are less populated and/or farther away from the business districts are more likely to receive subsidies. In addition, these users experience larger increases in their time windows.
7. For dynamic implementation of the incentive program, it is more beneficial to separate the budget for the morning and evening peak hours and distribute the available budget among re-optimization intervals.

There are at least three directions to extend the work in this paper. First, the fare paid by riders and the compensation received by drivers depend on their personal information (the valuations of time and distance), which we assumed could be estimated from their historical information. However, this information can also be directly solicited from the participants upon registration. In this case, it is essential to design a mechanism that ensures incentive compatibility, i.e., the individuals cannot benefit by not being truthful in reporting their private information (i.e., gaming the system). For future work, we consider designing a mechanism to determine the fare paid/received by riders/drivers based on their marginal contributions to social welfare (i.e., fairness), and also guarantees incentive compatibility (i.e., truthfulness). Moreover, this paper assumes the simplest form of ride-matching problem, i.e., the one-to-one ride-matching problem. As a result of this assumption, we observe that the IR incentive does not contribute much to improving social welfare. However, in case of allowing multiple riders or the possibility of transfers between vehicles, we might observe significant impacts by both types of incentives. Finally, this paper considers a simple linear model to incorporate the effect of subsidies on a user's utility. It is interesting to consider more complicated forms of subsidization in ridesharing, as Fang, Huang, and Wierman (2020) shows that the multi-threshold subsidy programs are more effective in practice. Additionally, participants' valuations of expanding their time windows from right or left (leaving later or earlier, respectively) could be different, especially for commuter trips.

Appendix A Table of Notations

Table A. 1. List of notations.

Notation		Definition
Sets:		
$N = \{n\}$		the set of all users
$R = \{r\}$		the set of all riders
$D = \{d\}$		the set of all drivers
$N' = \{n'\}$	$N'' = \{n''\}$	the set of trips in the morning/evening
$R' = \{r'\}$	$R'' = \{r''\}$	the set of rider trips in the morning/evening
$D' = \{d'\}$	$D'' = \{d''\}$	the set of driver trips in the morning/evening
N'''		the set of users with trips both in the morning and evening
$R' = \{r'\}$		the set of riders with trips both in the morning and evening
$D' = \{d'\}$		the set of drivers with trips both in the morning and evening
$A' = \{a'\}$	$A'' = \{a''\}$	the set of potential morning/evening trip matches
$V = \{v\}$		the set of nodes in the min-cost max flow network
$E = \{e\}$		the set of edges in the min-cost max flow network
Parameters:		
$\mathrm{I}(n')$	$\mathrm{I}(n'')$	the origin station of user n in the morning/evening
$\mathrm{J}(n')$	$\mathrm{J}(n'')$	the destination station of user n in the morning/evening
$\mathrm{T}(n')$	$\mathrm{T}(n'')$	the desired earliest departure time of user n in the morning/evening
$\mathrm{Q}(n')$	$\mathrm{Q}(n'')$	the desired latest arrival time of user n in the morning/evening
$\mathrm{F}(n)$		the valuation of every minute of travel time for user n
$\mathrm{H}(n)$		the valuation of every mile driven for user n
Γ		the maximum time window extension offered to a participant
B		the total budget available for monetary incentives
$\mathrm{U}(r'\|d')$	$\mathrm{U}(r''\|d'')$	the valuation of rider r if matched to driver d in the morning/evening
$\mathrm{U}(d'\|r')$	$\mathrm{U}(d''\|r'')$	the valuation of driver d if matched to rider r in the morning/evening
$\mathrm{W}(e)$		the original gain of pair of trips e
$\overline{\gamma}(d'\|r')$	$\overline{\gamma}(d''\|r'')$	the optimal BA incentive of driver d if matched with rider r in the morning/evening
$\overline{\gamma}(r'\|d')$	$\overline{\gamma}(r''\|d'')$	the optimal BA incentive of rider r if matched with driver d in the morning/evening
$\overline{\lambda}(d', r')$	$\overline{\lambda}(r'', d'')$	the optimal IR incentive of pair $(d', r')/(r'', d'')$ if matched in the morning/evening
$\overline{t}(d'\|r')$	$\overline{t}(d''\|r'')$	the optimal trip start time of driver d if matched with rider r in the morning/evening
$\overline{t}(r'\|d')$	$\overline{t}(r''\|d'')$	the optimal trip start time of rider r if matched with driver d in the morning/evening
$\Psi(d', r')$	$\Psi(r'', d'')$	the optimal subsidy allocated to the pair $(d', r')/(r'', d'')$ if matched in the morning/evening
s	t	the source/target node in the min-cost max flow network
ϵ		an infinitesimal positive, real number
$\mathrm{C}(e)$		the difference between the adjusted gain and subsidy of pair of trips e
$\mathrm{C}_{\max}$		the largest value of C
OPT		the optimal objective of the problem in (7)
β		the Lagrange multiplier
$\mathrm{C}_{\beta}(e)$		the coefficient of pair of trips e in LR subproblem for a given β
$\alpha(e)$		the weight of edge e in the Gasoline Lemma
ϕ		a flat tax rate for budget balanced variant of TIP
π		a rate for dividing the budget between the morning and evening periods
Functions:		
G		the min-cost flow network of a ridesharing system with ride-back guarantee
G'		the graph containing the cycles between two flows S_l and S_h
G''		the min-cost flow network of an infeasible flow S
$\delta^{+}(v)$	$\delta^{-}(v)$	the in-going and out-going edges of node v in G

Tables:		
τ		the shortest-path travel times matrix for the stations
ρ		the shortest-path driving distances matrix for the stations
Variables:		
$x(e)$		1 if the driver and rider trip in e are matched, and 0 otherwise
$t(n')$	$t(n'')$	the morning/evening trip start time of user n
$q(n')$	$q(n'')$	the morning/evening trip end time of user n
$\gamma^-(n')$	$\gamma^-(n'')$	the morning/evening time extension of the earliest departure time for user n
$\gamma^+(n')$	$\gamma^+(n'')$	the morning/evening time extension of the latest arrival time for user n
$\gamma(n')$	$\gamma(n'')$	the BA incentive for the trip of user n in the morning/evening
$\lambda(e)$		the IR incentive for the pair of trips e
$w(e)$		the adjusted gain of pair of trips e including the subsidies
$S_h = \{e\}$	$S_l = \{e\}$	A budget-feasible/infeasible flow that maximizes the LR subproblem
$\tilde{S} = \{e\}$		A near-optimal solution obtained from Algorithm 2

Appendix B Gasoline Lemma

Lemma 1. *((Lin and Kernighan 1973)) Given a sequence of* k *real numbers* $\alpha_0, \dots, \alpha_{k-1}$ *such that* $\sum_{j=0}^{k-1} \alpha_j = 0$*, there is an index* $i \in \{0, \dots, k-1\}$ *such that, for any* $0 \le h \le k-1$,

$$\sum_{j=i}^{i+h} \alpha_{j(\bmod k)} \ge 0, \quad \forall\ h \in \{0, \dots, k-1\}.$$

Proof. Let $i' \in \{0, \dots, k-1\}$ be the index for which $\sum_{j=0}^{i\prime} \alpha_j$ is minimum and let i be $(i'+1)(\bmod k)$. Thus, we have:

$$\sum_{j=i}^{i+h} \alpha_{j(\bmod k)} = \sum_{j=0}^{i+h} \alpha_{j(\bmod k)} - \sum_{j=0}^{i\prime} \alpha_{j(\bmod k)} \ge 0, \quad \forall\ h \in \{0, \dots, k-1\}.$$

Appendix C Proof of Proposition 1

Let us consider different possible scenarios of subsequence Y and show that in each scenario we can obtain a feasible flow by removing a set of edges from S. Let us denote the first and last edges of subsequence Y as $e_1 = (i_1, j_1)$ and $e_2 = (i_2, j_2)$, respectively. For any node in subsequence Y which is not incident to these 2 edges, it is easy to show that the flow conservation constraint is satisfied in $S = (S_h \cup Y_f) \backslash Y_b$. However, this is not the case for i_1 and j_2.

From the assumptions on Y and the procedure described in the proof of the Gasoline Lemma, we infer that e_1 is one of the following six scenarios: (i) (d', r'), $(ii)(r', d')$, $(iii)(d'', r'')$, $(iv)(r'', d'')$, $(v)(r', r'')$, and $(vi)(r'', r')$. In scenario (i), by adding (d', r') to S_h the flow conservation will be violated if d' is matched to another rider trip in S_h. Therefore, we must remove edge (d', r') from S. Moreover, if $r \in R'''$, we may have to remove $(r'', d_s(r''))$ from S in the worst case. Note that $d_s(r'')$ is the driver trip matched with rider r in the evening. In scenario (ii), removing (d', r') from S_h will violate flow conservation for node r'' in S if $r \in R'''$. Therefore, we may have to remove edge $(r'', d_s(r''))$ in the worst case. Using the same line of reasoning as in scenario (ii), we can show that we have to remove (r'', d'') for the fourth scenario, and $(r'', d_s(r''))$ for the fifth and sixth scenarios from S in the worst-case. Note that scenario (iii) does not require removing any edges from S. It is worth mentioning that we may also need to remove some of the auxiliary edges in S which involve the source and target nodes, but those edges have no effect on the objective function value.

From the fact that Y is the longest subseqence in X, we conclude that e_2 cannot be of type (d', r'), (r'', d''), and (r'', r'), because their following edges in X will have to be included in Y. As such, e_2 can only follow one of the following scenarios: $(i)(r', d')$, $(ii)(d'', r'')$, and $(iii)(r', r'')$. In the first case, we do not need to remove any edge from S other than the auxiliary edges. In the next two scenarios, however, we have to remove $(d_s(r'), r')$ if $r \in R'''$.

Let r''_1 and r_2'' be two rider trips in the evening for $r_1 \in R'''$ and $r_2 \in R'''$. Now, let us consider the following possible cases for Y based on the scenarios for $e_1 = (i_1, j_1)$ and $e_2 = (i_2, j_2)$:

8. $i_1 \neq r''_1$ and $j_2 \neq r''_2$: In this case, the property of Y in (14) implies that:

$$\sum_{e \in Y} \alpha(e) = \sum_{e \in Y_f} \mathsf{C}_{\beta^*}(e) - \sum_{e \in Y_b} \mathsf{C}_{\beta^*}(e) \geq 0\,.$$

   By adding $\mathsf{C}_{\beta^*}(S_h)$ and $\beta^* \, \Psi(S)$ to both sides of the inequality, we have:

$$\mathsf{C}(S) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S)\,.$$

   Also, based on the scenarios discussed above, at most two edges will be removed from S to get a feasible solution for the min-cost max flow in graph G'' in this case. Therefore, $\mathsf{C}(\tilde{S}) \geq \mathsf{C}(S) - 2\,\mathsf{C}_{\max}$, because $\tilde{S}$ is a feasible solution that maximizes the costs in graph G''. Thus, we conclude that $\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}$.

9. $i_1 = r_1''$ and $j_2 \neq r_2''$: In this case, the property in (14) implies that:

$$\mathsf{C}(S) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - \max\{\mathsf{C}(d'_l(r_1'), r_1'),\ \mathsf{C}(d'_h(r_1'), r_1'),\ 0\}\,.$$

   Also, based on the scenarios discussed above, at most one edge will be removed in this case which yields $\mathsf{C}(\tilde{S}) \geq \mathsf{C}(S) - \mathsf{C}_{\max}$. Thus, we again conclude that $\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}$.

10. $i_1 \neq r''_1$ and $j_2 = r''_2$: In this case, the property in (14) implies that:

$$\mathsf{C}(S) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) + \max\{\mathsf{C}(d'_l(r_2'), r_2'),\ \mathsf{C}(d'_h(r_2'), r_2'),\ 0\}\,.$$

   Based on the scenarios discussed above, at most two edges will be removed due to i_1 and edge $(d_s(r_2'), r_2')$ due to j_2 in this case. As a result, we have $\mathsf{C}(\tilde{S}) \geq \mathsf{C}(S) - 2\,\mathsf{C}_{\max} - \mathsf{C}(d_s(r_2'), r_2')$. Note that $d_s(r_2')$ is either $d_h(r_2')$ or $d_l(r_2')$. Thus, we again conclude that $\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}$.

11. $i_1 = r_1''$ and $j_2 = r_2''$: In this case, the property in (14) implies that:

$$\begin{aligned}\mathsf{C}(S) \geq\ & \mathsf{C}_{\beta^*}(S_h) - \beta^* \, \Psi(S) - \max\{\mathsf{C}(d'_l(r_1'), r_1'),\ \mathsf{C}(d'_h(r_1'), r_1'),\ 0\} + \\ & \max\{\mathsf{C}(d'_l(r_2'), r_2'),\ \mathsf{C}(d'_h(r_2'), r_2'),\ 0\}\,.\end{aligned}$$

Based on the scenarios discussed above, at most one edge will be removed due to i_1 and edge $(d_s(r_2'), r_2')$ due to j_2 in this case. Therefore, $\mathsf{C}(\tilde{S}) \geq \mathsf{C}(S) - 2\,\mathsf{C}_{\max} - \mathsf{C}(d_s(r_2'), r_2')$. Using the same reasoning as in the last two cases, we can again conclude that $\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}$.

In all the cases above, we have that $\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}$. Let us rewrite this inequlaity as:

$$\mathsf{C}(\tilde{S}) \geq \mathsf{C}_{\beta^*}(S_h) + \beta^* \, \mathsf{B} - \beta^* \, \mathsf{B} + \beta^* \, \Psi(S) - 2\,\mathsf{C}_{\max}\,. \tag{19}$$

We know that $\mathsf{C}_{\beta^*}(S_h) + \beta^* \, \mathsf{B}$ is a solution to the LR subproblem, and thus, we have:

$$\mathsf{C}_{\beta^*}(S_h) + \beta^* \, \mathsf{B} \geq \mathsf{OPT}\,. \tag{20}$$

Also, from the fact that Y is the longest subsequence in X for which $\Psi(S) \leq \mathsf{B}$, we know that there always exist an edge $e_3 \in X \backslash Y$ such that $\Psi(e_3) + \Psi(S) > \mathsf{B}$. Now, consider the two following cases for edge e_3:

12. $\mathsf{C}_{\beta^*}(e_3) \geq 0$: In this case, $\mathsf{C}(e_3) \geq \beta^* \Psi(e_3)$ which yields $-\beta^* \Psi(e_3) \geq -\mathsf{C}(e_3) \geq -\mathsf{C}_{\max}$.
13. $\mathsf{C}_{\beta^*}(e_3) < 0$: Due to the fact that e_3 is a part the min-cost max flow solution in graph G with costs set as $-\mathsf{C}_{\beta^*}$, this case can happen only if there exists an edge e_4 such that $\mathsf{C}_{\beta^*}(e_3) + \mathsf{C}_{\beta^*}(e_4) \geq 0$ which implies that $\mathsf{C}(e_4) \geq \beta^* \Psi(e_3)$. Thus, again, we have $-\beta^* \Psi(e_3) \geq -\mathsf{C}(e_4) \geq -\mathsf{C}_{\max}$.

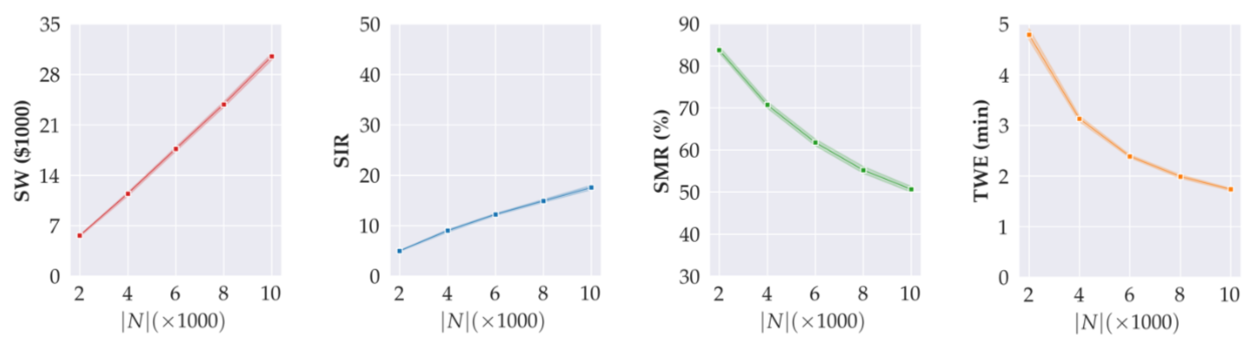
From the cases above, we conclude that

$$-\beta^* \mathsf{B} + \beta^* \Psi(S) \geq -\mathsf{C}_{\max}. \tag{21}$$

Combining the inequalities in (19), (20), and (21) yields $\mathsf{C}(\tilde{S}) \geq \mathsf{OPT} - 3\,\mathsf{C}_{\max}$ and the result follows.

Appendix D Sensitivity Analysis (Cont'd)

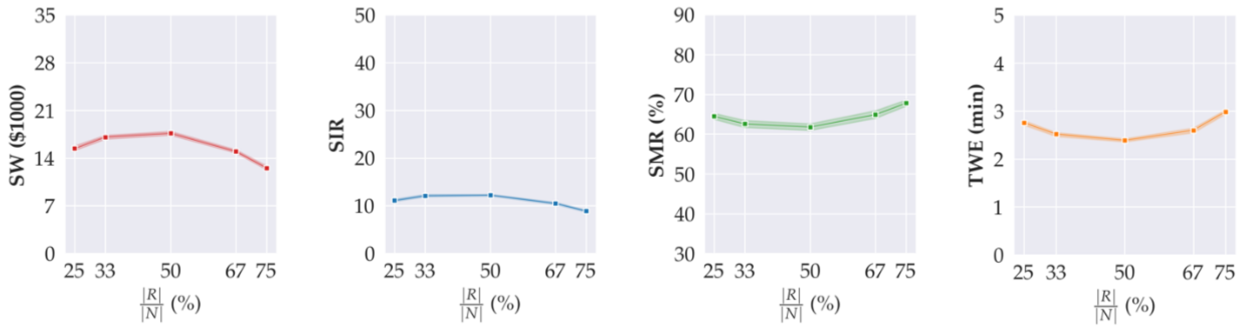
In the base scenario, we assume that the total number of participants in the ridesharing system during the morning and evening peak hours is 6000. In order to study the impact of the number of participants on different performance metrics, we change its value to 2000, 4000, 6000, 8000, and 10,000. The results are presented in Figure D.1. Figure D.1(a) suggests that the social welfare of the system increases significantly as the penetration rate of the system increases. Also, the rate of increase in these figures is linear, which is not surprising because the system is not saturated and adding more users increases the possibility of matching, and thereby, increasing the system's social welfare. The upward trend of the subsidy impact rate in Figure D. 1(b) indicates that the added value to the system per one dollar spent on subsidy significantly increases with the number of participants. This is due to the fact that as the number of participants grows, the spatio-temporal proximity of trips increases, which results in (i) fewer users requiring the BA incentive to get matched, and (ii) a smaller amount of the BA incentive for those who need it (see Figure D.1(d)). Finally D.1 (c) and D.1(d) suggest that at least 50% of the matches are subsidized while the average time window extension for those that received the BA incentive does not exceed 5 minutes. The higher spatio-temporal proximity of trips that follows from a higher penetration rate allows the system to subsidize fewer rides and the participants to experience less deviation from their preferred time windows.



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure D. 1: Impact of number of participants

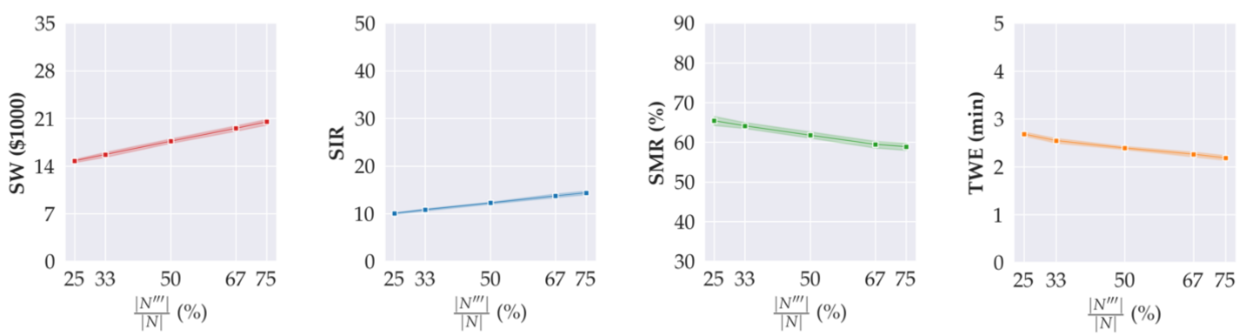
The number of riders and drivers are assumed to be the same in the base scenario. Next, we investigate the impact of changing the percentage of riders from 50 to 25, 33, 67 and 75, respectively. Figure D.2 displays the results of these scenarios for different performance metrics. Figure D.2(a) clearly shows that the social welfare decreases when the percentage of riders diverges from 50%, especially when the percentage of riders is greater than the percentage of drivers. This is partly due to the fact that we are implementing a one-to-one system where a single driver carries at most a single rider in each peak period, and therefore the best results are obtained when there is a balance between the number of riders and drivers. When we have fewer riders than drivers, more drivers are available to serve them, and therefore the percentage of matched riders will be higher than the case where we have more riders than drivers due to (1) more resources to match the riders, and (2) fewer riders to be served. This trend can also be partly explained by the assumption of ride-back guarantees, as 50% of riders in these scenarios register both their morning and evening trips in the system and will be served if and only if both of their trips are served by the drivers in the system. Since this type of requests are harder to satisfy, we expect that the matching rate and social welfare decrease with a higher rate for scenarios above 50%. On top of having a smaller social welfare, Figure D.2(c) shows that a higher percentage of matches need to be subsidized when the percentage of riders is higher than 50. This is why the subsidy impact rates are lower for these scenarios compared to the base scenario, as shown in Figure D.2(b).



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure D. 2: Impact of percentage of riders

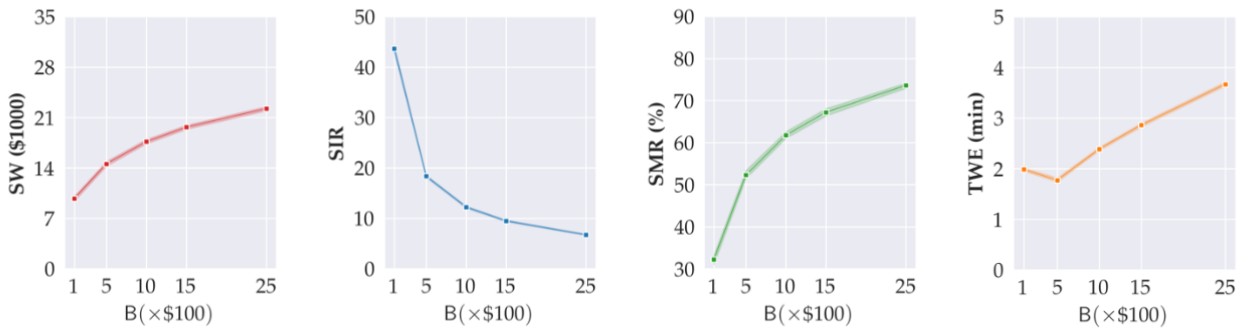
The base scenario assumes that 50% of users register both their morning and evening trips in the system. In this part, we investigate the impact of having lower percentages (25% and 33%) or higher percentages (67%, 75%) of participants be present in both the morning and evening peak hours. Figure D.3 demonstrates how different performance metrics change as this parameter increases from 25% to 75%. Given a fixed number of participants, increasing the value of this parameter clearly increases the number of users (both riders and drivers) in the morning and in the evening. Thus, the number of matches and hence the social welfare increases as shown in Figure D.3(a). Also, the upward trend of the subsidy impact rate in Figure D.3(b) implies that the proposed incentive program is more beneficial when users register both trips in the system. The downward trend of subsidized matched users in Figure D.3(c) originates from the fact that with a higher number of riders and drivers in the morning or evening, the likelihood of finding a match without the help of the BA incentive increases, because of the higher spatio-temporal proximity between trips. Higher spatio-temporal proximity between trips further explains the linearly decreasing trend in D.3(d), as users will be required to expand their time window less when this parameter increases.



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure D. 3: Impact of percentage of ride-back trips

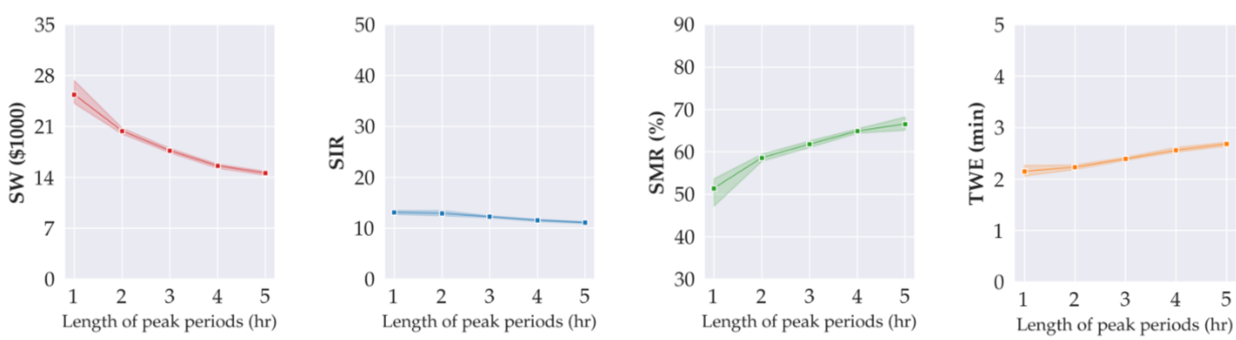
Figure D.4 demonstrates the results of changing the available budget for subsidy from \$1,000 in the base scenario to \$100, \$500, \$1500 and \$2,500. Figure D.4(a) shows that the social welfare increases sublinearly as we increase the value of B. This is consistent with our earlier results in Section 5.6, where we observed that social welfare grows sublinearly until it converges to the maximum possible social welfare. More interestingly, the same trend is also observed in Figure D.4(c). Moreover, the subsidy impact rate in Figure D.4(b) has the reverse trend: we can initially make a huge impact by investing a small amount of budget, but the rate of return-on-investment diminishes as we increase the budget. However, as we try to incentivize higher number of matches, we have to invest more, and the margin of profit gets smaller. Finally, we reach a point that either no further matches can become feasible with the help of the BA incentive or the required budget becomes higher than its subsequent added social welfare. Figure D.4(d) shows an interesting result where the time window expansion decreases slightly as we increase the budget from \$100 to \$500 and then increases from \$500 to \$2500. One possible explanation would be that from \$100 to \$500, the rate of increase in the number of subsidized matches is very high (more than 20%) while the amount of increase in time windows for those users is not that much higher, and thus, the average time window expansion slightly decreases.



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure D. 4: Impact of total budget

Finally, Figure D.5 displays the impact of the length of peak hour periods in the morning and the evening on the performance metrics. In the base scenario, we assume that the length of both peak hour periods is 3 hours. Here, we consider tighter periods (1 and 2 hours) and wider periods (4 and 5 hours). Obviously, increasing the length of the peak hour periods causes the trips to spread over a larger horizon, which results in reducing the temporal proximity of trips. This is the main reason behind the descending trend in the social welfare presented in Figures D.5(a). Also, for the same reason, more participants need the BA incentive (see Figure D.5(c)), and the magnitude of allocated incentive per user increases as shown in Figure D.5(d). Moreover, since trips become more temporally sparse, the savings due to sharing rides decreases, which consequently decreases the impact of each dollar spent on subsidy, as shown in Figure D.5(b).



(a) Social welfare (b) Subsidy impact rate (c) Subsidized matches rate (d) Time window extension

Figure D. 5: Impact of the morning and evening peak hours' lengths

References

Agatz, Niels, Alan L Erera, Martin WP Savelsbergh, and Xing Wang. 2011. "Dynamic Ride-Sharing: A Simulation Study in Metro Atlanta." *Procedia-Social and Behavioral Sciences* 17: 532–50.

Agatz, Niels, Alan Erera, Martin Savelsbergh, and Xing Wang. 2012. "Optimization for Dynamic Ride-Sharing: A Review." *European Journal of Operational Research* 223 (2): 295–303.

Ahuja, Ravindra K, Thomas L Magnanti, and James B Orlin. 1988. "Network Flows."

Alonso-Mora, Javier, Samitha Samaranayake, Alex Wallar, Emilio Frazzoli, and Daniela Rus. 2017. "On-Demand High-Capacity Ride-Sharing via Dynamic Trip-Vehicle Assignment." *Proceedings of the National Academy of Sciences* 114 (3): 462–67.

Amey, Andrew. 2011. "A Proposed Methodology for Estimating Rideshare Viability Within an Organization, Applied to the Mit Community." In *TRB Annual Meeting Procediings*, 1–16.

Baldacci, Roberto, Vittorio Maniezzo, and Aristide Mingozzi. 2004. "An Exact Method for the Car Pooling Problem Based on Lagrangean Column Generation." *Operations Research* 52 (3): 422–39.

Becker, Henrik, Milos Balac, Francesco Ciari, and Kay W Axhausen. 2020. "Assessing the Welfare Impacts of Shared Mobility and Mobility as a Service (MaaS)." *Transportation Research Part A: Policy and Practice* 131: 228–43.

Berger, André, Vincenzo Bonifaci, Fabrizio Grandoni, and Guido Schäfer. 2011. "Budgeted Matching and Budgeted Matroid Intersection via the Gasoline Puzzle." *Mathematical Programming* 128 (1): 355–72.

Bian, Zheyong, and Xiang Liu. 2019. "Mechanism Design for First-Mile Ridesharing Based on Personalized Requirements Part i: Theoretical Analysis in Generalized Scenarios." *Transportation Research Part B: Methodological* 120: 147–71.

Brownstone, David, and Thomas F Golob. 1992. "The Effectiveness of Ridesharing Incentives: Discrete-Choice Models of Commuting in Southern California." *Regional Science and Urban Economics* 22 (1): 5–24.

Chan, Nelson D, and Susan A Shaheen. 2012. "Ridesharing in North America: Past, Present, and Future." *Transport Reviews* 32 (1): 93–112.

Chen, Wenyi, Martijn Mes, Marco Schutten, and Job Quint. 2019. "A Ride-Sharing Problem with Meeting Points and Return Restrictions." *Transportation Science* 53 (2): 401–26.

Deakin, Elizabeth, Karen Trapenberg Frick, and Kevin M Shively. 2010. "Markets for Dynamic Ridesharing? Case of Berkeley, California." *Transportation Research Record* 2187 (1): 131–37.

Fang, Zhixuan, Longbo Huang, and Adam Wierman. 2020. "Loyalty Programs in the Sharing Economy: Optimality and Competition." *Performance Evaluation*, 102105.

Fisher, Marshall L. 1981. "The Lagrangian Relaxation Method for Solving Integer Programming Problems." *Management Science* 27 (1): 1–18.

Furuhata, Masabumi, Maged Dessouky, Fernando Ordóñez, Marc-Etienne Brunet, Xiaoqing Wang, and Sven Koenig. 2013. "Ridesharing: The State-of-the-Art and Future Directions." *Transportation Research Part B: Methodological* 57: 28–46.

Gallo, Giorgio, and Claudio Sodini. 1978. "Extreme Points and Adjacency Relationship in the Flow Polytope." *Calcolo* 15 (3): 277–88.

Hasan, Mohd Hafiz, Pascal Van Hentenryck, and Antoine Legrain. 2020. "The Commute Trip-Sharing Problem." *Transportation Science* 54 (6): 1640–75.

Hawkins, Andrew J. 2019. "Uber and Lyft Finally Admit They're Making Traffic Congestion Worse in Cities - the Verge." https://www.theverge.com/2019/8/6/20756945/uber-lyft-tnc-vmt-traffic-congestion-study-fehr-peers.

Holzhauser, Michael, Sven O Krumke, and Clemens Thielen. 2016. "Budget-Constrained Minimum Cost Flows." *Journal of Combinatorial Optimization* 31 (4): 1720–45.

Johnson, Donald B. 1975. "Finding All the Elementary Circuits of a Directed Graph." *SIAM Journal on Computing* 4 (1): 77–84.

Lee, Alan, and Martin Savelsbergh. 2015. "Dynamic Ridesharing: Is There a Role for Dedicated Drivers?" *Transportation Research Part B: Methodological* 81: 483–97.

Li, Ruijie, Yu Nie, and Xiaobo Liu. 2020. "Pricing Carpool Rides Based on Schedule Displacement." *Transportation Science* 54 (4): 1134–52.

Lin, Shen, and Brian W Kernighan. 1973. "An Effective Heuristic Algorithm for the Traveling-Salesman Problem." *Operations Research* 21 (2): 498–516.

Lloret-Batlle, Roger, Neda Masoud, and Daisik Nam. 2017. "Peer-to-Peer Ridesharing with Ride-Back on High-Occupancy-Vehicle Lanes: Toward a Practical Alternative Mode for Daily Commuting." *Transportation Research Record* 2668 (1): 21–28.

Long, Jiancheng, Weimin Tan, WY Szeto, and Yao Li. 2018. "Ride-Sharing with Travel Time Uncertainty." *Transportation Research Part B: Methodological* 118: 143–71.

Luo, Qi, Romesh Saigal, Zhibin Chen, and Yafeng Yin. 2019. "Accelerating the Adoption of Automated Vehicles by Subsidies: A Dynamic Games Approach." *Transportation Research Part B: Methodological* 129: 226–43.

Ma, Shuo, Yu Zheng, and Ouri Wolfson. 2013. "T-Share: A Large-Scale Dynamic Taxi Ridesharing Service." In *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, 410–21. IEEE.

Masoud, Neda, and R Jayakrishnan. 2017a. "A Decomposition Algorithm to Solve the Multi-Hop Peer-to-Peer Ride-Matching Problem." *Transportation Research Part B: Methodological* 99: 1–29.

———. 2017b. "A Real-Time Algorithm to Solve the Peer-to-Peer Ride-Matching Problem in a Flexible Ridesharing System." *Transportation Research Part B: Methodological* 106: 218–36.

Megiddo, Nimrod. 1978. "Combinatorial Optimization with Rational Objective Functions." In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, 1–12.

Najmi, Ali, David Rey, and Taha H Rashidi. 2017. "Novel Dynamic Formulations for Real-Time Ride-Sharing Systems." *Transportation Research Part E: Logistics and Transportation Review* 108: 122–40.

Nam, Daisik, Dingtong Yang, Sunghi An, Jiangbo Gabriel Yu, R Jayakrishnan, and Neda Masoud. 2018. "Designing a Transit-Feeder System Using Multiple Sustainable Modes: Peer-to-Peer (P2p) Ridesharing, Bike Sharing, and Walking." *Transportation Research Record*, 0361198118799031.

Nourinejad, Mehdi, and Matthew J Roorda. 2016. "Agent Based Model for Dynamic Ridesharing." *Transportation Research Part C: Emerging Technologies* 64: 117–32.

Qian, Xinwu, Wenbo Zhang, Satish V Ukkusuri, and Chao Yang. 2017. "Optimal Assignment and Incentive Design in the Taxi Group Ride Problem." *Transportation Research Part B: Methodological* 103: 208–26.

Regue, Robert, Neda Masoud, and Will Recker. 2016. "Car2work: A Shared Mobility Concept to Connect Commuters with Workplaces." *Transportation Research Record: Journal of the Transportation Research Board* 2542: 102–10.

Schneider, Benjamin. 2018. "New Study of Global Traffic Reveals That Traffic Is Bad - Bloomberg." https://www.bloomberg.com/news/articles/2018-02-07/new-study-of-global-traffic-reveals-that-traffic-is-bad.

Stiglic, Mitja, Niels Agatz, Martin Savelsbergh, and Mirko Gradisar. 2015. "The Benefits of Meeting Points in Ride-Sharing Systems." *Transportation Research Part B: Methodological* 82: 36–53.

———. 2016. "Making Dynamic Ride-Sharing Work: The Impact of Driver and Rider Flexibility." *Transportation Research Part E: Logistics and Transportation Review* 91: 190–207.

Su, Qing, and Liren Zhou. 2012. "Parking Management, Financial Subsidies to Alternatives to Drive Alone and Commute Mode Choices in Seattle." *Regional Science and Urban Economics* 42 (1-2): 88–97.

Tafreshian, Amirmahdi, and Neda Masoud. 2020a. "Trip-Based Graph Partitioning in Dynamic Ridesharing." *Transportation Research Part C: Emerging Technologies* 114: 532–53.

———. 2020b. "Using Subsidies to Stabilize Peer-to-Peer Ridesharing Markets with Role Assignment." *Transportation Research Part C: Emerging Technologies* 120: 102770.

Tafreshian, Amirmahdi, Neda Masoud, and Yafeng Yin. 2020. "Ride-Matching for Peer-to-Peer Ridesharing: A Review and Future Directions." *Service Science*, to appear.

Wang, Xiaolei, Hai Yang, and Daoli Zhu. 2018. "Driver-Rider Cost-Sharing Strategies and Equilibria in a Ridesharing Program." *Transportation Science* 52 (4): 868–81.

Wang, Xiaoqing, Maged Dessouky, and Fernando Ordonez. 2016. "A Pickup and Delivery Problem for Ridesharing Considering Congestion." *Transportation Letters* 8 (5): 259–69.

Wang, Xing, Niels Agatz, and Alan Erera. 2017. "Stable Matching for Dynamic Ride-Sharing Systems." *Transportation Science*.

Xiong, Chenfeng, Mehrdad Shahabi, Jun Zhao, Yafeng Yin, Xuesong Zhou, and Lei Zhang. 2020. "An Integrated and Personalized Traveler Information and Incentive Scheme for Energy Efficient Mobility Systems." *Transportation Research Part C: Emerging Technologies* 113: 57–73.

Xu, Zhengtian, Yafeng Yin, and Jieping Ye. 2020. "On the Supply Curve of Ride-Hailing Systems." *Transportation Research Part B: Methodological* 132: 29–43.

Zhao, Daozhi, and Mingyang Chen. 2019. "Ex-Ante Versus Ex-Post Destination Information Model for on-Demand Service Ride-Sharing Platform." *Annals of Operations Research* 279 (1-2): 301–41.