Vehicle Rebalancing in A Shared Micromobility System with Rider Crowdsourcing

Ziliang Jin[a], Yulan Wang[a], Yun Fong Lim[b], Kai Pan[a,*], Zuo-Jun Max Shen[c,d]

[a]Faculty of Business, The Hong Kong Polytechnic University, Kowloon, Hong Kong

[b]Lee Kong Chian School of Business, Singapore Management University, Singapore

[c]College of Engineering, University of California, Berkeley, California 94720, USA

[d]Faculty of Engineering and Faculty of Business and Economics, The University of Hong Kong, Hong Kong

[*]Corresponding author

Contact: ziliang-lms.jin@connect.polyu.hk (ZJ), yulan.wang@polyu.edu.hk (YW), yflim@smu.edu.sg (YFL), kai.pan@polyu.edu.hk (KP), maxshen@berkeley.edu (Z-JMS)

Problem definition: Shared micromobility vehicles provide an eco-friendly form of short-distance travel within an urban area. Since customers pick up and drop off vehicles in any service region at any time, such convenience often leads to a severe imbalance between vehicle supply and demand in different service regions. To overcome this, a micromobility operator can crowdsource individual riders with reward incentives in addition to engaging a third-party logistics provider (3PL) to relocate the vehicles.

Methodology/results: We construct a time-space network with multiple service regions and formulate a two-stage stochastic mixed-integer program considering uncertain customer demands. In the first stage, the operator decides the initial vehicle allocation for the regions, whereas in the second stage, he determines subsequent vehicle relocation across the regions over an operational horizon. We develop an efficient solution approach that incorporates scenario-based and time-based decomposition techniques. Our approach outperforms a commercial solver in solution quality and computational time for solving large-scale problem instances based on real data.

Managerial implications: The budgets for acquiring vehicles and for rider crowdsourcing significantly impact the vehicle initial allocation and subsequent relocation. Introducing rider crowdsourcing in addition to the 3PL can significantly increase profit, reduce demand loss, and improve the vehicle utilization rate of the system without affecting any existing commitment with the 3PL. The 3PL is more efficient for mass relocation than rider crowdsourcing, while the latter is more efficient in handling sporadic relocation needs. To serve a region, the 3PL often relocates vehicles in batches from faraway, low-demand regions around peak hours of a day, whereas rider crowdsourcing relocates a few vehicles each time from neighboring regions throughout the day. Furthermore, rider crowdsourcing relocates more vehicles under a unimodal customer arrival pattern than a bimodal pattern, whereas the reverse holds for the 3PL.

Key words: Shared Micromobility; Crowdsourcing; Allocation and Relocation; Two-stage Stochastic Mixed-integer Programming; Decomposition Algorithm

1. Introduction

A *shared micromobility system* consists of lightweight vehicles such as bikes, e-bikes, e-scooters, or e-mopeds. It offers an eco-friendly form of short-distance travel such as last-mile transportation and helps alleviate city traffic congestion, achieving a sustainable urban transportation system (Simlett and Møller 2020, McKinsey 2021). During the COVID-19 pandemic, the shared micromobility system becomes even more popular as some people avoid taking crowded public transportation. In the U.S., around 60% of car trips are within 8 kilometers, making the shared micromobility system an excellent alternative. It is estimated that the micromobility market in the U.S. alone will reach 200 to 300 billion dollars by 2030 (McKinsey 2019).

A shared micromobility system differs from ride hailing systems mainly because of the following three features. First, the shared micromobility system is asset heavy and its operator often bears high investment and operational costs of its physical assets (e.g., bikes) (Hasija et al. 2020). In a competitive market, these high costs cannot be covered by simply increasing rental prices because customers can easily switch to other mobility services such as public transportation or car sharing. Thus, it is important for the operator to properly determine a total number of vehicles and allocate them to each service region to satisfy demand and generate profitability. Second, as there is no self-regulated owner for each micromobility vehicle, shared micromobility services are not provided on an on-demand basis like ride hailing because drivers do not ride a micromobility vehicle to search for customers. Third, customers of the shared micromobility system can pick up and drop off vehicles in any service region at any time. Such convenience often leads to a severe imbalance between the vehicle supply and demand in different regions. As shown in our detailed analyses in Appendix B, such imbalance results in an oversupply of vehicles with few pick-ups in some service regions and insufficient vehicles in others. This can substantially undermine the system's operational efficiency, service quality, and profitability. As a result, the shared micromobility system requires efficient relocation of its vehicles across the service regions.

The above three features draw special attention to the *initial allocation* and *subsequent relocation* of vehicles to help the shared micromobility system achieve economic viability and sustainability. In contrast to the ride hailing systems that adopt instruments such as increasing drivers' salaries or dynamic pricing to control the vehicle supply and demand, the operator of the shared micromobility system can first allocate and then relocate vehicles to balance their supply and demand for each service region. This allows the operator to boost profitability given the low profit margin of micromobility services.

Currently, many shared micromobility operators outsource the vehicle relocation to third-party logistics providers (3PLs). Each 3PL rebalances micromobility vehicles across service regions in batches of vehicles primarily by trucks to enjoy the economies of scale (Dell'Amico et al. 2014). Such 3PL relocation is often executed according to a certain schedule, for example, in the early morning or late night, and may not meet all vehicle-relocation requirements. Each request for the 3PL relocation incurs a fixed setup cost and no additional fees are required if the 3PL takes multiple time periods to fulfill this request. If the 3PL relocates many vehicles for each request, then its relocation cost *per vehicle* is low. In fact, due to the high setup cost incurred in each request, it is too costly to frequently engage the 3PL for vehicle relocation across the service regions to rebalance the supply and demand.

To overcome this challenge, some shared micromobility firms adopt a crowdsourcing strategy to hire *individual riders* to relocate vehicles. Each crowdsourced rider relocates one vehicle at a time and receives a reward from the operator upon completing the task. In addition to the financial rewards, the riders are motivated because the relocation trips may be close to the riders' original travel routes or they may simply use this opportunity to exercise. For example, Mobike rewards riders with coupons or cash to relocate bikes (Horwitz 2017) and Bird pays riders to charge and relocate e-scooters (Bird 2020). Such rider crowdsourcing can help relocate vehicles in a more flexible manner. In contrast to the 3PL relocation, rider crowdsourcing offers more flexibility by providing sporadic relocation — relocating a small number of vehicles that is not justifiable for the 3PL. In this situation, the relocation cost *per vehicle* by rider crowdsourcing is lower, resulting in a lower total relocation cost. The vehicles relocated by crowdsourcing can be used to serve customer demands across service regions in each period, improving the revenue and reducing demand loss. Thus, the resulting benefit from rider crowdsourcing can be significant.

To the best of our understanding, our paper is the first to study the effectiveness and implications of combining rider crowdsourcing with a 3PL's service for an integrated allocation and relocation problem of a shared micromobility system over multiple periods under demand uncertainty. In particular, we are interested in the following research questions: How should a shared micromobility operator integrate the vehicle allocation and relocation decisions to match his finite supply with uncertain demand such that his profit is maximized? Given a limited budget for acquiring vehicles, how does the budget affect the initial vehicle allocation? How does this initial allocation affect the subsequent vehicle relocation across the regions? Similarly, how does a budget for crowdsourcing individual riders affect the vehicle allocation and relocation? What is the best strategy for vehicle relocation (i.e., rider crowdsourcing only, 3PL only, both of them, or none of them)? Under each relocation strategy, are there any interesting temporal or spatial features of the shared micromobility system? How do system parameters affect the system performance?

In this paper, we consider an operator that provides a fleet of micromobility vehicles to satisfy customer demands in a service area over an operational horizon with multiple periods. Due to government regulations, the entire service area is divided into multiple regions with different vehicle allocation capacities (He et al. 2017, Qi et al. 2018). The customer demands in each period are uncertain. At the start of the horizon, the operator first decides the initial vehicle allocation for the different service regions without knowing the actual demands in each period. Subsequently, in each period, after the demands are realized, the operator determines and executes the vehicle relocation across the regions (using rider crowdsourcing and the 3PL) to match the vehicles with customer demands. Unsatisfied demands in each period are lost. The operator's objective is to maximize his expected profit over the operational horizon.

We make the following contributions in this paper:

(i) We formulate the operator's *integrated vehicle allocation and relocation problem* as a two-stage stochastic mixed-integer program (Birge and Louveaux 2011) on a time-space network. In the first stage, we decide the initial vehicle allocation for the service regions, whereas in the second stage we determine the subsequent vehicle relocation across the regions over the horizon. The vehicle relocation is modeled as recourse decisions after the demands are realized in each period.

(ii) We develop an efficient algorithmic approach that incorporates scenario-based and time-based (temporal) decomposition ideas to obtain high-quality solutions to the problem. Our numerical experiments based on data collected from Citi Bike (2021) in New York City (NYC) suggest that our approach yields better solutions in a much shorter time than a commercial solver.

(iii) We obtain the following managerial insights for the micromobility operator. The budget for acquiring vehicles and the budget for rider crowdsourcing have significant impact on the initial vehicle allocation and the subsequent vehicle relocation. We find that introducing rider crowdsourcing in addition to the 3PL can significantly increase profit, reduce demand loss, and improve the vehicle utilization rate of the system without affecting the existing commitment with the 3PL. The 3PL is more efficient in mass relocation than rider crowdsourcing, whereas the latter is more efficient in handling sporadic relocation needs. The 3PL often conducts relocation around peak hours of a day by moving vehicles in batches from faraway, low-demand regions. In contrast, rider crowdsourcing is engaged throughout the day to relocate a few vehicles each time from neighboring regions. Furthermore, rider crowdsourcing relocates more vehicles under a unimodal customer arrival pattern than a bimodal pattern, whereas the reverse holds for the 3PL.

After reviewing the related literature in Section 2, we formulate the problem in Section 3. Section 4 studies structural results for the two-region case. Section 5 develops a solution approach for the general case with multiple regions. Section 6 performs extensive numerical experiments to examine the two vehicle-relocation methods based on the data from Citi Bike (2021). Section 7 concludes the paper. All proofs are presented in Online Supplement.

2. Literature Review

Our paper is closely related to the Operations Management (OM) studies on shared mobility, a component of smart cities (Qi and Shen 2019, Mak 2022). Most papers on shared mobility focus on vehicle-sharing systems, where people rent or hail cars from individual car owners or drivers for trips (e.g., Nair and Miller-Hooks 2014, Boyacı et al. 2015, and Feng et al. 2021). He et al. (2017) consider a service region design problem by incorporating fleet operations and the customer adoption rate for electric vehicles under uncertainty. Chang et al. (2017) optimize the car fleet location, size, and type via integer programming by considering a carbon emission constraint. Lu et al. (2018) optimize vehicle allocation under demand uncertainty and address the impact of one-way and round-way trips on the system's profit and service quality.

Shared micromobility, a type of shared mobility, further brings challenges and opportunities to OM researchers (Hasija et al. 2020). Recent work by Kabra et al. (2020) empirically investigates the impact of bike accessibility and availability on bike-share ridership by using a structural demand model. Our work is more related to Shu et al. (2013). Specifically, they build linear programming models with proportionality constraints to optimize the bicycle flows when the initial allocation of bicycles is given at each dock station and to examine the impact of bicycle operations on the dock size. However, the decisions on initial allocation and fleet operations are inherently interdependent. Therefore, we consider an integrated allocation and relocation model in which an operator decides the initial allocation and subsequent relocation of vehicles.

Our paper also contributes to the literature on fleet operations of shared mobility. Existing studies extend the broader scope of inventory redistribution or transshipment (Benjaafar et al. 2022), which has been widely investigated especially for the optimal stocking policy (Tagaras and Cohen 1992, Grahovac and Chakravarty 2001). Different approaches have been proposed to facilitate fleet operations for vehicle-sharing systems. Nair and Miller-Hooks (2011) minimize the cost of fleet redistribution by mixed-integer programming with chance constraints to ensure a service level. Lu et al. (2018) implement dynamic vehicle relocation via a rolling-horizon method. He et al. (2020) derive an optimal relocation policy for a two-region system and use a linear decision rule to solve a distributionally robust model for a multi-region system.

Our work is more related to micromobility vehicle operations such as fleet operations of bicycle-sharing systems. Dell'Amico et al. (2014) model bike relocation by trucks as a capacitated pickup-and-delivery vehicle routing problem and use a branch-and-cut algorithm to find solutions. Freund et al. (2020) consider truck- and trike-based rebalancing approaches for a bike-sharing system, and formulate the underlying routing problems using integer programming to minimize expected customer dissatisfaction. Li and Liu (2021) consider a bike rebalancing problem over a single period with deterministic demand and derive a static operation policy that employs trucks and users to relocate bikes. Fu et al. (2022) build a two-stage robust model to optimize bike allocation in the first stage and fleet operations in the second stage. The above studies often ignore the system operator's capacity design (e.g., bike allocation), which is more critical in shared micromobility than in vehicle-sharing as mentioned in Section 1. Furthermore, the above papers mostly consider bike relocation in batches and the corresponding route design problem. In contrast, our paper considers different relocation strategies using rider crowdsourcing and a 3PL under demand uncertainty.

Finally, our paper is related to crowdsourcing in the OM literature. Crowdsourcing is widely studied in last-mile delivery operations, where independent car drivers (Qi et al. 2018, Fatehi and Wagner 2022) or cyclists (Kafle et al. 2017) are incentivized to serve customers with fast and on-time delivery. Unlike goods that exit the system once they are delivered, shared micromobility vehicles stay in the system after they are relocated. Such a difference inspires further studies to establish foundations for using crowdsourcing in micromobility vehicle relocation. According to a survey by Singla et al. (2015), customers are willing to alter their routes to help relocate bikes given monetary incentives. Fricker and Gast (2016) analyze the steady-state performance of a bicycle-sharing system and show that incentivizing users to the least-loaded station leads to fewer problematic stations. He et al. (2021) study a two-sided matching platform, such as a free-float bike-sharing platform, that determines both parking space allocation and the design of incentive instruments (e.g., price and rewards to customers) to control the supply in each region. Huang et al. (2021) find that a sparse structure of a transshipment network can guide bike repositioning by crowdsourced volunteers and help reduce workload rebalancing without much demand loss. Stokkink and Geroliminis (2021) consider a station-based one-way car-sharing system and incentivize customers to perform car relocation. Cheng et al. (2021) consider a free-floating bike-sharing system and propose a bidding model to incentivize users to relocate bikes. Unlike the above studies, our paper offers new managerial insights by incorporating rider crowdsourcing and a 3PL's service into an integrated micromobility vehicle allocation and relocation model under different temporal demand patterns.

3. Problem Formulation

Consider an operator that provides shared micromobility service for a set of regions $\mathcal{V} = \{1, 2, \ldots, V\}$ in each period $t \in \mathcal{T} = \{0, 1, \ldots, T-1\}$. We first study the movements of micromobility vehicles across the service regions and then construct an optimization model.

3.1. Vehicle Movements: A Time-Space Network

We model the vehicle movements across regions and periods as flows in a time-space network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is a set of nodes and $\mathcal{A}$ is a set of directed arcs on the network as shown in Figure 1. Each node $n_{is} \in \mathcal{N}$ represents service region $i \in \mathcal{V}$ in period $s \in \mathcal{T}$. The *flow* on the directed arc $(n_{is}, n_{jt}) \in \mathcal{A}$ with $s \leq t-1$ represents the number of vehicles moving from node n_{is} to node n_{jt}. At the start of period 0, each node n_{i0} is allocated with an initial number of vehicles x_i for $i \in \mathcal{V}$. At the end of the operational horizon (period $T-1$), the number of vehicles in each node $n_{i,T-1}$ is denoted as $\widetilde{x}_i$ for $i \in \mathcal{V}$.

We assume that a rider takes $l_{ij}(\geq 1)$ periods to ride a micromobility vehicle from region i to region j for $i, j \in \mathcal{V}$. For convenience, we set $l_{ii} = 1$ for $i \in \mathcal{V}$. In contrast, we assume the 3PL takes a fixed number of periods $l_{\mathrm{r}}(\geq 1)$ to relocate vehicles from region i to region j. This assumption is reasonable for a relatively compact geographical area in which the regions in $\mathcal{V}$ are near each other. To ensure that a micromobility vehicle rider has enough time to move from region i to region j before the end of the operational horizon, we consider the demands from region i to region j that occur only in period $t \in \mathcal{T}(l_{ij}) = \{0, 1, \ldots, T - l_{ij} - 1\}$. Similarly, to ensure that the 3PL can finish all the relocations before the end of the operational horizon, we consider 3PL relocations that begin only in period $t \in \mathcal{T}(l_{\mathrm{r}}) = \{0, 1, \ldots, T - l_{\mathrm{r}} - 1\}$.

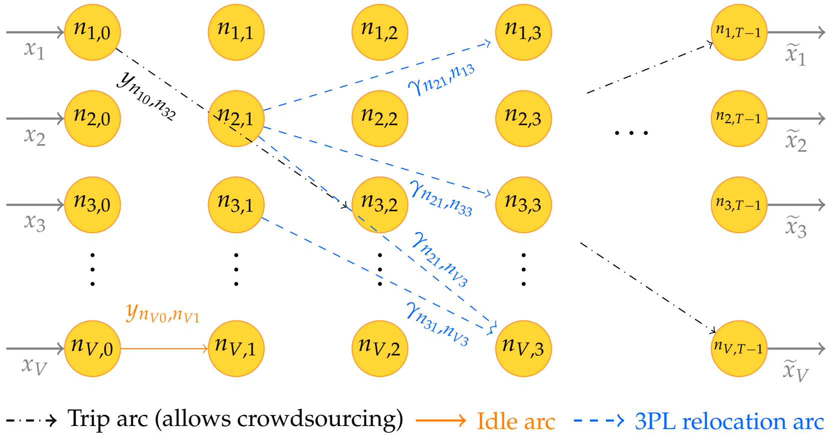


Figure 1 Time-space Network $\mathcal{G}$

Based on how micromobility vehicles are moved between two nodes, there are three types of arcs in $\mathcal{A}$ in the time-space network: (i) *Trip arcs*: The flow of each trip arc $(n_{is}, n_{jt}) \in \mathcal{A}^{\mathrm{t}}$ represents the number of trips from region i in period $s \in \mathcal{T}(l_{ij})$ to region j in period $t = s + l_{ij}$, for $i \neq j, i, j \in \mathcal{V}$. Each arc in $\mathcal{A}^{\mathrm{t}}$ may contain two types of trips: one corresponds to the customer demand and the other is induced by crowdsourced riders. (ii) *Idle arcs*: The flow of each idle arc $(n_{jt}, n_{j,t+1}) \in \mathcal{A}^{\mathrm{i}}$, represents the number of idling vehicles in region $j \in \mathcal{V}$ from period $t \in \mathcal{T} \backslash \{T-1\}$ to period $t+1$. (iii) *3PL relocation arcs*: The flow of each 3PL relocation arc $(n_{is}, n_{j,s+l_{\mathrm{r}}}) \in \mathcal{A}^{\mathrm{r}}$ represents the number of vehicles relocated by the 3PL from region i in period $s \in \mathcal{T}(l_{\mathrm{r}})$ to region j, for $i \neq j, i, j \in \mathcal{V}$. Define $\mathcal{I} = \{\mathrm{t}, \mathrm{i}, \mathrm{r}\}$. Thus, we have $\mathcal{A} = \cup_{\forall e \in \mathcal{I}} \mathcal{A}^{e}$ and $\mathcal{A}^{e_1} \cap \mathcal{A}^{e_2} = \emptyset$, for $e_1 \neq e_2$, $e_1, e_2 \in \mathcal{I}$.

Figure 2 illustrates the sequence of events regarding vehicle movements and demand arrivals at each node $i \in \mathcal{V}$: (i) At the start of period $t \in \mathcal{T} \setminus \{T-1\}$, the numbers of idling vehicles in region i from period $t-1$ and incoming vehicles from region $j \in \mathcal{V} \setminus \{i\}$ are realized. (ii) Customer demands arrive. (iii) Knowing the realized demands, the operator determines and executes the numbers of vehicles relocated by crowdsourced riders and the 3PL. (iv) Any demand loss is observed and the system is updated for period $t+1$.

We formulate the problem as a *two-stage stochastic program*. In the first stage, we decide the initial number of vehicles allocated to each region. In the second stage, we decide the vehicle relocation in each period for the entire operational horizon as a recourse.

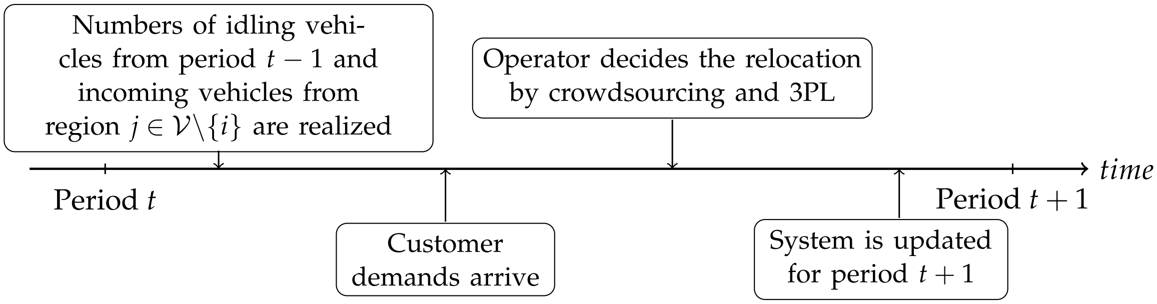

Numbers of idling vehicles from period $t-1$ and incoming vehicles from region $j \in \mathcal{V}\backslash\{i\}$ are realized
Operator decides the relocation by crowdsourcing and 3PL
time
Period t
Period $t+1$
Customer demands arrive
System is updated for period $t+1$

Figure 2 The sequence of events

3.2. Mathematical Formulation

The shared micromobility operator is endowed with a budget that can afford at most N micromobility vehicles. In the first stage, the operator makes the vehicle allocation decisions $\mathbf{x} = (x_i,\ i \in \mathcal{V})^\top$ so that $\sum_{j\in\mathcal{V}} x_j \le N$. A cost $c_j(\ge 0)$ is incurred for allocating a micromobility vehicle to region $j \in \mathcal{V}$. Due to government regulations on the parking space, we have $x_j \le B_j$ for each region $j \in \mathcal{V}$.

In the second stage, time-dependent demands arrive at each region $i \in \mathcal{V}$ at the start of each period $t \in \mathcal{T}\backslash\{T-1\}$. For each arc $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$, let $\lambda_{n_{it},n_{j,t+l_{ij}}}$ denote the demand (number of customers) from region i in period t to region j with travel time l_{ij}. The operator receives a revenue $R(\ge 0)$ for serving a customer per period. If the number of vehicles available in region i cannot meet the demands of the region in period t, the unsatisfied demand for $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$, denoted by $\eta_{n_{it},n_{j,t+l_{ij}}}$, is lost and the operator pays a penalty cost $C_{\mathrm{p}}(\ge 0)$ per customer lost.

Facing uncertain demands, the operator makes his initial vehicle allocation decisions and subsequent relocation decisions to maximize his expected profit. The profit equals the total revenue minus the total cost. The total cost includes the initial vehicle allocation cost $\sum_{j\in\mathcal{V}} c_j x_j$ in the first stage and the demand-loss penalty and vehicle relocation cost in the second stage. Given the initial vehicle allocation $\mathbf{x}$, let $\Theta(\mathbf{x})$ denote the optimal expected net cost of the second stage (i.e., the total cost minus the total revenue of the second stage). Thus, maximizing the operator's expected profit is equivalent to minimizing the summation of $\sum_{j\in\mathcal{V}} c_j x_j$ and $\Theta(\mathbf{x})$. The operator optimizes his vehicle allocation decisions by solving the following problem:

$$\Gamma = \min_{\mathbf{x}} \sum_{j\in\mathcal{V}} c_j x_j + \Theta(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{X} = \left\{ \mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \;\middle|\; x_j \le B_j,\ j \in \mathcal{V},\ \sum_{j\in\mathcal{V}} x_j \le N \right\}, \tag{$\mathcal{M}$}$$

where $\Theta(\mathbf{x})$ is realized in the second stage as the operator optimizes the vehicle relocation using rider crowdsourcing and the 3PL. Problem ($\mathcal{M}$) is an *integrated vehicle allocation and relocation problem*. We discuss the vehicle relocation problem in the second stage in detail below.

Vehicle Relocation by Crowdsourced Riders: For each trip arc $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$, we introduce a continuous decision variable $\Lambda_{n_{it},n_{j,t+l_{ij}}}$ to represent the number of crowdsourced riders from region i in period t to region j. The operator provides an incentive $\phi_{n_{it},n_{j,t+l_{ij}}}$ to motivate the riders to relocate the vehicles along arc $(n_{it}, n_{j,t+l_{ij}})$. We assume rider crowdsourcing follows the seminal law of diminishing returns (Shephard and Färe 1974): The marginal reward for the crowdsourced riders increases with the number of crowdsourced riders. Specifically, we define a concave increasing incentive function $g(\cdot)$ such that

$$\Lambda_{n_{it},n_{j,t+l_{ij}}} = g\left(\phi_{n_{it},n_{j,t+l_{ij}}}\right) = \bar{\Lambda}_{ij} \times \left(1 - e^{-\beta_{ij}\phi_{n_{it},n_{j,t+l_{ij}}}}\right),\ t \in \mathcal{T}(l_{ij}),\ i \neq j,\ i,j \in \mathcal{V}, \tag{1a}$$

where $\bar{\Lambda}_{ij} \geq 0$ represents the maximum number of riders that can be crowdsourced in region i to travel to region j, and $\beta_{ij} \geq 0$ denotes the rate of diminishing return on rewards. Both $\bar{\Lambda}_{ij}$ and β_{ij} can be estimated from historical data. We adopt the nonlinear crowdsourcing supply curve in (1a) to align with the literature where the law of diminishing returns has been widely applied in many fields (Cannan 1892, Rosenberg 1992, Averbakh and Berman 1996, Chung et al. 1997, Srinivasan and Moorman 2005, Homburg et al. 2015, Ravichandran et al. 2017).

As mentioned in Section 1, the shared micromobility system is asset heavy with a low profit margin. To reflect the practice that the operator is often financially constrained, we require that the total incentive to crowdsource riders is capped by a budget $B_{\mathrm{c}}(\geq 0)$:

$$\sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}, j \neq i} \sum_{t \in \mathcal{T}(l_{ij})} \phi_{n_{it},n_{j,t+l_{ij}}} \leq B_{\mathrm{c}}. \tag{1b}$$

For each arc $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{i}} \cup \mathcal{A}^{\mathrm{t}}$, let $y_{n_{it},n_{j,t+l_{ij}}}$ denote its realized flow. We have

$$y_{n_{it},n_{j,t+l_{ij}}} = \lambda_{n_{it},n_{j,t+l_{ij}}} + \Lambda_{n_{it},n_{j,t+l_{ij}}} - \eta_{n_{it},n_{j,t+l_{ij}}},\ t \in \mathcal{T}(l_{ij}),\ i \neq j,\ i,j \in \mathcal{V}. \tag{1c}$$

Vehicle Relocation by the 3PL: For each period $t \in \mathcal{T}(l_{\mathrm{r}})$, we define a decision variable z_t such that $z_t = 1$ if the 3PL is requested to *start* relocating vehicles from period t onward, and $z_t = 0$ otherwise. In practice, due to the 3PL's capacity limitation, the operator can only request the 3PL for vehicle relocation for no more than $\bar{z}$ times over the entire operational horizon:

$$\sum_{t \in \mathcal{T}(l_{\mathrm{r}})} z_t \leq \bar{z}. \tag{2a}$$

For each request, the 3PL charges the operator a fixed fee $C_{\mathrm{r}}(\geq 0)$ to relocate vehicles across regions. No additional fees are required if it takes multiple periods to relocate the vehicles. Define a binary variable $\widetilde{z}_t$ to represent the 3PL's status: $\widetilde{z}_t = 1$ if the 3PL provides its service in period t, and $\widetilde{z}_t = 0$ otherwise. Since the 3PL provides service only if the operator has requested it, we have

$$z_t - \widetilde{z}_t \leq 0,\ t \in \mathcal{T}(l_{\mathrm{r}}). \tag{2b}$$

EXAMPLE 1. This example illustrates the difference between z_t and $\tilde{z}_t$. Suppose the 3PL is first requested to start relocating vehicles from period s_1 (i.e., $z_{s_1} = 1$) as shown in Figure 3. The 3PL provides its service for two consecutive periods (i.e., $\widetilde{z}_{s_1} = \widetilde{z}_{s_1+1} = 1$, corresponding to filled circles). The 3PL then stops its service from period $s_1 + 2$ until period $s_2 - 1$ (i.e., $\widetilde{z}_{s_1+2} = \cdots = \widetilde{z}_{s_2-1} = 0$, corresponding to empty circles). The 3PL is requested again to start relocation from period s_2 (i.e., $z_{s_2} = 1$) and provides its service for three consecutive periods (i.e., $\widetilde{z}_{s_2} = \widetilde{z}_{s_2+1} = \widetilde{z}_{s_2+2} = 1$).

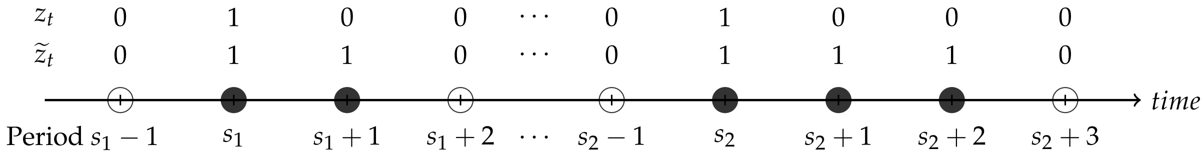


Figure 3 Service Status of the 3PL

As illustrated in Example 1, z_t and $\widetilde{z}_t$ can have different values for each period t, while their logical relationship is described by constraints (2b)–(2d) and (2g). Note that the variable $\widetilde{z}_t$ is necessary to indicate whether the 3PL provides its service (i.e., its service status) in period t, while the variable z_t indicates whether the 3PL is requested to start relocating vehicles from period t.

If the operator makes a request for the 3PL's service in period t (i.e., $z_t = 1$), then it implies that no 3PL relocation service is provided in period $t - 1$ (i.e., $\widetilde{z}_{t-1} = 0$). We have

$$z_t + \widetilde{z}_{t-1} - 1 \leq 0,\ t \in \mathcal{T}(l_\mathrm{r}) \backslash \{0\}. \tag{2c}$$

If a 3PL relocation service is provided in period t (i.e., $\widetilde{z}_t = 1$) but not in period $t - 1$ (i.e., $\widetilde{z}_{t-1} = 0$), then clearly the relocation service is initiated in period t (i.e., $z_t = 1$). This can be represented by

$$z_t - \widetilde{z}_t + \widetilde{z}_{t-1} \geq 0,\ t \in \mathcal{T}(l_\mathrm{r}) \backslash \{0\}. \tag{2d}$$

Due to its capacity limitation, the 3PL can serve for at most $\bar{\bar{z}}$ periods in any interval of l_f periods:

$$\sum_{i=t}^{t+l_\mathrm{f}-1} \widetilde{z}_i \leq \bar{\bar{z}},\ t \in \{0, 1, \ldots, T - l_\mathrm{r} - l_\mathrm{f}\}. \tag{2e}$$

For each arc $(n_{it}, n_{j,t+l_\mathrm{r}}) \in \mathcal{A}^\mathrm{r}$, define a continuous decision variable $\gamma_{n_{it}, n_{j,t+l_\mathrm{r}}}$ to represent the number of vehicles relocated by the 3PL from region i in period t to region j. The total number of relocated vehicles in period t is constrained between a lower limit $\underline{q}$ and an upper limit $\bar{q}$:

$$\underline{q}\widetilde{z}_t \leq \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}, j \neq i} \gamma_{n_{it}, n_{j,t+l_\mathrm{r}}} \leq \bar{q}\widetilde{z}_t,\ t \in \mathcal{T}(l_\mathrm{r}). \tag{2f}$$

At the start of period 0, we assume that the 3PL is requested if it provides relocation service:

$$z_0 - \widetilde{z}_0 \geq 0. \tag{2g}$$

Finally, we restrict $\widetilde{z}_t$ to be binary:

$$\widetilde{z}_t \in \{0, 1\},\ t \in \mathcal{T}(l_\mathrm{r}), \tag{2h}$$

which, together with constraints (2b)–(2d) and (2g), guarantee z_t to be binary for $t \in \mathcal{T}(l_\mathrm{r})$.

Flow Balance: For each node $n_{it} \in \mathcal{N}, i \in \mathcal{V}$, the number of vehicles flowing into this node should equal the number of vehicles flowing out from this node:

$$\left[y_{n_{it},n_{i,t+1}} + \sum_{j\in\mathcal{V},j\neq i}\left(y_{n_{it},n_{j,t+l_{ij}}} + \gamma_{n_{it},n_{j,t+l_\mathrm{r}}}\right)\right] - \left[y_{n_{i,t-1},n_{it}} + \sum_{j\in\mathcal{V},j\neq i}\left(y_{n_{j,t-l_{ji}},n_{it}} + \gamma_{n_{j,t-l_\mathrm{r}},n_{it}}\right)\right]$$
$$= \begin{cases} x_i, & \text{if } t = 0; \\ 0, & \text{if } t = 1,\ldots,T-2; \\ -\widetilde{x}_i, & \text{if } t = T-1. \end{cases} \tag{3}$$

Note that the above constraints may contain some undefined arcs (n_{is}, n_{jt}) with $s < 0$ or $t \geq T$ for $i \neq j,\ i,j \in \mathcal{V}$ (see Figure 19 in Appendix C.1). For consistency, we set the realized flow $y_{n_{is},n_{jt}} = 0$ and the number of vehicles relocated by the 3PL $\gamma_{n_{is},n_{jt}} = 0$ on such arcs.

When crowdsourced riders relocate a small number of vehicles not justifiable for the 3PL, the relocation cost *per vehicle* is low by the incentive curve in (1a), leading to a low total relocation cost. However, as the number of crowdsourced riders increases, the relocation cost per vehicle becomes larger because of the concave increasing function in (1a). In contrast, if the 3PL relocates many vehicles for each request by the operator, then its relocation cost per vehicle is low. Thus, the relocation cost per vehicle by rider crowdsourcing can be smaller or larger than that by the 3PL. The operator needs to optimize the vehicle relocation in each period by considering the relocation cost per vehicle of each method.

In the second stage, given an initial vehicle allocation $\mathbf{x}$, the operator relocates the vehicles to maximize his expected profit subject to uncertain demands $\lambda_{n_{it},n_{j,t+l_{ij}}}$. We assume that there is a finite support for the joint distribution of the uncertain demands in all the service regions across all the periods. We use a set $\mathcal{K}$ to include scenarios of uncertain demands in all the service regions across all the periods. Each scenario $k \in \mathcal{K}$ has a probability $p^k (\geq 0)$ with $\sum_{k\in\mathcal{K}} p_k = 1$. For each scenario $k \in \mathcal{K}$, the operator makes recourse decisions for all the periods in $\mathcal{T}$. We reuse the above notation for the second-stage problem and add a superscript k to each decision variable for each scenario k. We use bold symbols to represent vectors. For example, $\widetilde{\mathbf{x}}^k = (\widetilde{x}_j^k,\ j \in \mathcal{V})^\top$ (see Appendix C.2 for details). For each $k \in \mathcal{K}$, let $\mathbf{Y}^k = (\widetilde{\mathbf{x}}^k, \boldsymbol{\eta}^k, \mathbf{y}^k, \boldsymbol{\gamma}^k, \boldsymbol{\Lambda}^k, \boldsymbol{\phi}^k, \mathbf{z}^k, \widetilde{\mathbf{z}}^k)$ denote all the decision variables, $\mathbf{Y}^{k*}$ be the optimal solution, and $\mathcal{Y}(\mathbf{x}, \boldsymbol{\lambda}^k) := \{\mathbf{Y}^k \in \mathbb{R}_+^{(|\mathcal{V}|+4|\mathcal{A}^\mathrm{t}|+|\mathcal{A}^\mathrm{r}|+2|\mathcal{T}(l_\mathrm{r})|)} \mid (1a)-(1c), (2a)-(2h), (3)\}$ denote the feasible region. Given the initial vehicle allocation $\mathbf{x}$, the expected net cost $\Theta(\mathbf{x})$ in the second stage can be determined by solving the following network flow optimization problem:

$$\Theta(\mathbf{x}) = \min_{\mathbf{Y}^k \in \mathcal{Y}(\mathbf{x},\boldsymbol{\lambda}^k),\ k\in\mathcal{K}} \left\{ \sum_{k\in\mathcal{K}} p^k \left[\sum_{a\in\mathcal{A}^\mathrm{t}} \left(C_\mathrm{p}\eta_a^k + \phi_a^k - R l_a \left(y_a^k - \Lambda_a^k \right) \right) + \sum_{t\in\mathcal{T}(l_\mathrm{r})} C_\mathrm{r} z_t^k \right] \right\}. \tag{$\mathcal{P}$}$$

Note that the flows on the idle arcs do not incur any costs and the flows corresponding to the 3PL and crowdsourcing relocation do not generate revenue.

4. Analysis of Two-region Case

We consider a special case of problem ($\mathcal{M}$) with two service regions. To obtain structural results, we focus on a scenario in $\mathcal{K}$ with demands $\lambda_{n_{it},n_{3-i,t+L}}$. We assume that the initial vehicle allocation $\mathbf{x} = (x_1, x_2)^\top$ is given so that we can study the optimal relocation strategy in detail. We focus on the following problem setting: (i) The travel time and the 3PL relocation time between the two regions are a constant L. (ii) To relocate vehicles from region i in period t to the other region by crowdsourcing, the total reward provided increases linearly with the number of crowdsourced riders: $\phi_{n_{it},n_{3-i,t+L}} = \alpha\Lambda_{n_{it},n_{3-i,t+L}} + b$, where $\alpha(\geq 0)$ represents a variable cost of crowdsourcing an additional rider and $b(\geq 0)$ represents a fixed cost of crowdsourcing. (iii) We set $\underline{q} = 0$ and $\bar{q} = \max\left\{\lambda_{n_{it},n_{3-i,t+L}} \;\middle|\; i \in \{1,2\}, t \in \{0,1,\dots,T-L-1\}\right\}$. Problem ($\mathcal{P}$) becomes

$$\Theta(x_1, x_2) = \min_{(\boldsymbol{\Lambda},\mathbf{y},\boldsymbol{\eta},\mathbf{z},\boldsymbol{\gamma})\in\mathcal{Y}(x_1,x_2,\boldsymbol{\lambda})} \left\{ \sum_{t=0}^{T-L-1} \left[\sum_{i=1}^{2} \Big(C_{\mathrm{p}}\eta_{n_{it},n_{3-i,t+L}} + \alpha\Lambda_{n_{it},n_{3-i,t+L}} + b \right.\right.$$
$$\left.\left. - RL\left(y_{n_{it},n_{3-i,t+L}} - \Lambda_{n_{it},n_{3-i,t+L}}\right)\Big) + C_{\mathrm{r}} z_t \right] \right\}, \tag{4}$$

where $\mathcal{Y}(x_1, x_2, \boldsymbol{\lambda})$ is described in Appendix D.1. Note that there are no incoming vehicles for any region in the first L periods and no demands in the last L periods. Thus, it is sufficient to consider periods $t \in \{0, 1, \dots, T-L-1\}$ in the objective.

PROPOSITION 1. *Given any $i \in \{1,2\}$ and $t \in \{0,1,\dots,T-L-1\}$, the optimal relocation strategy satisfies the following three conditions: (a) $\Lambda^*_{n_{it},n_{3-i,t+L}}\gamma^*_{n_{it},n_{3-i,t+L}} = 0$; (b) $C_r \geq \alpha\Lambda^*_{n_{it},n_{3-i,t+L}}$; and (c) if $z^*_t = 1$ and $\gamma^*_{n_{it},n_{3-i,t+L}} > 0$, then $C_r \leq \alpha\gamma^*_{n_{it},n_{3-i,t+L}}$.*

Part (a) of Proposition 1 shows that under the optimal relocation strategy, the 3PL and crowdsourcing are not used simultaneously in any region i and period t. Part (b) shows that if crowdsourcing is used (i.e., $\Lambda^*_{n_{it},n_{3-i,t+L}} > 0$), then its relocation cost per vehicle is less than that by the 3PL (i.e., $\alpha \leq C_{\mathrm{r}}/\Lambda^*_{n_{it},n_{3-i,t+L}}$). Similarly, Part (c) shows that if the 3PL is used in period t (i.e., $z^*_t = 1$ and $\gamma^*_{n_{it},n_{3-i,t+L}} > 0$), then its relocation cost per vehicle is less than that by crowdsourcing (i.e., $C_{\mathrm{r}}/\gamma^*_{n_{it},n_{3-i,t+L}} \leq \alpha$). Such insights imply that the relocation cost per vehicle of each method serves as a critical instrument for the operator to choose a method for each arc. Section 6 considers a more practical setting in industry by solving the general problem ($\mathcal{M}$). We first introduce a solution approach to problem ($\mathcal{M}$) in the following section.

5. Solution Approach for Multiple Regions

To solve problem ($\mathcal{M}$), we first propose a piecewise-linear approximation of the crowdsourcing incentive function $g(\cdot)$ in (1a) and then develop a decomposition-based algorithm to solve it.

5.1. Piecewise-Linear Approximation of the Incentive Function

In (1a), given any scenario $k \in \mathcal{K}$ and arc $(n_{it}, n_{i,t+l_{ij}}) \in \mathcal{A}^t$, the number of crowdsourced riders is $\Lambda^k_{n_{it},n_{i,t+l_{ij}}} = g(\phi^k_{n_{it},n_{i,t+l_{ij}}}) = \bar{\Lambda}_{ij}(1 - e^{-\beta_{ij}\phi^k_{n_{it},n_{j,t+l_{ij}}}})$, which is a concave increasing function of the reward amount $\phi^k_{n_{it},n_{i,t+l_{ij}}}$. To make problem ($\mathcal{M}$) tractable, we approximate $g(\cdot)$ with a piecewise-linear function (see Figure 4) such that the nonlinear constraints (1a) can be replaced by linear constraints. Since the approximation is applicable for all i, j, t, and k, we drop the superscript and subscripts such that $\Lambda = g(\phi) = \bar{\Lambda}(1 - e^{-\beta\phi})$, where $\phi \in [0, \infty)$.

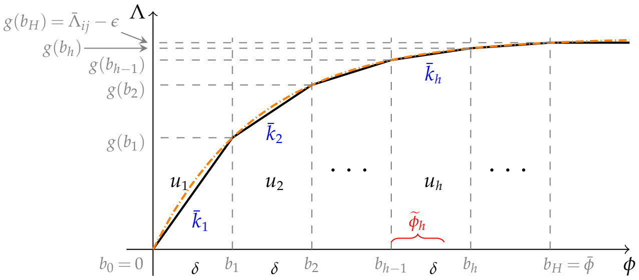


Figure 4 Piecewise-Linear Approximation of the Incentive Function

Given an arbitrarily small $\epsilon(> 0)$, define $\bar{\phi} = g^{-1}(\bar{\Lambda} - \epsilon)$. We divide the interval $[0, \bar{\phi}]$ into H segments, each with length $\delta = \bar{\phi}/H$. For any $h \in \mathcal{H} = \{1, 2, \ldots, H\}$, we approximate $g(\cdot)$ in the hth segment $[(h-1)\delta, h\delta]$ (denoted as $[b_{h-1}, b_h]$ in Figure 4) with a linear function passing through points $((h-1)\delta, g((h-1)\delta))$ and $(h\delta, g(h\delta))$. Let $\bar{k}_h = (g(h\delta) - g((h-1)\delta))/\delta$ denote the slope of this linear function. For ϕ in $[\bar{\phi}, \infty)$, we approximate $g(\cdot)$ with a constant $\bar{\Lambda} - \epsilon$. For each $h \in \mathcal{H}$, the number of crowdsourced riders Λ linearly increases with ϕ, while the slope $\bar{k}_h$ is decreasing in h and becomes 0 as ϕ reaches $\bar{\phi}$. Clearly, a larger H leads to a more accurate approximation.

For each $h \in \mathcal{H}$, define a binary decision variable u_h such that u_h = 1 if the optimal reward ϕ falls in the hth segment, and $u_h = 0$ otherwise. Since we minimize the objective of problem ($\mathcal{M}$) and no additional crowdsourced riders will be attracted when $\phi \geq \bar{\phi}$ (i.e., the optimal ϕ will not exceed $\bar{\phi}$), no binary variables are needed for $[\bar{\phi}, \infty)$. We thus restrict $\phi \in [0, \bar{\phi}]$. For each $h \in \mathcal{H}$, define a continuous decision variable $\widetilde{\phi}_h \in [0, \delta]$ such that the optimal reward $\phi = (h-1)\delta + \widetilde{\phi}_h$ and the optimal number of crowdsourced riders $\Lambda = g((h-1)\delta) + \bar{k}_h\widetilde{\phi}_h$ in the hth segment. Thus, the constraint $\Lambda = g(\phi)$ can be approximated by the following set of constraints:

$$\phi = \sum_{h=1}^{H} u_h \left((h-1)\delta + \widetilde{\phi}_h\right), \tag{5a}$$

$$\Lambda = \sum_{h=1}^{H} u_h \left(\sum_{m=1}^{h-1} \bar{k}_m\delta + \bar{k}_h\widetilde{\phi}_h\right), \tag{5b}$$

$$\sum_{h=1}^{H} u_h = 1, \tag{5c}$$

$$0 \leq \widetilde{\phi}_h \leq \delta,\ h \in \mathcal{H}, \tag{5d}$$

$$u_h \in \{0, 1\},\ h \in \mathcal{H}. \tag{5e}$$

Since u_h is binary, we can exactly reformulate the bilinear terms $u_h\widetilde{\phi}_h$ in (5a) and (5b) by replacing them with ω_h in the following linear constraints:

$$\phi = \sum_{h=1}^{H} \left((h-1)u_h\delta + \omega_h\right),\ \Lambda = \sum_{h=1}^{H} u_h \sum_{m=1}^{h-1} \bar{k}_m\delta + \sum_{h=1}^{H} \bar{k}_h\omega_h,$$

$$u_h\delta - \omega_h \geq 0,\ \widetilde{\phi}_h - \omega_h \geq 0,\ \delta - u_h\delta - \widetilde{\phi}_h + \omega_h \geq 0,\ \omega_h \geq 0,\ h \in \mathcal{H}. \tag{6}$$

Based on (5c)–(5e) and (6), we are ready to reformulate the relocation problem ($\mathcal{P}$) by approximating nonlinear constraints (1a) for each scenario $k \in \mathcal{K}$. For each arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$, we add a subscript a to the decision variables $\phi, u_h, \widetilde{\phi}_h$, and ω_h, and add subscripts ij to the parameters δ and $\bar{k}_h$ in (5c)–(5e) and (6). Given an initial allocation $\mathbf{x}$, problem ($\mathcal{P}$) can be decomposed into $|\mathcal{K}|$ independent problems $\mathcal{P}^k$, $k \in \mathcal{K}$. The net cost of $\mathcal{P}^k$ equals

$$\Psi\left(\widetilde{\mathbf{Y}}^k\right) = \sum_{a \in \mathcal{A}^{\mathrm{t}}} \left[C_{\mathrm{p}}\eta_a^k + \phi_a^k - Rl_a\left(y_a^k - \Lambda_a^k\right)\right] + \sum_{t \in \mathcal{T}(l_{\mathrm{r}})} C_{\mathrm{r}} z_t^k, \tag{7}$$

where $\widetilde{\mathbf{Y}}^k = (\widetilde{\mathbf{x}}^k, \boldsymbol{\eta}^k, \mathbf{y}^k, \gamma^k, \boldsymbol{\Lambda}^k, \mathbf{u}^k, \boldsymbol{\phi}^k, \widetilde{\boldsymbol{\phi}}^k, \mathbf{z}^k, \widetilde{\mathbf{z}}^k, \boldsymbol{\omega}^k) \in \mathbb{R}_+^{(|\mathcal{V}|+(4+3H)|\mathcal{A}^{\mathrm{t}}|+|\mathcal{A}^{\mathrm{r}}|+2|\mathcal{T}(l_{\mathrm{r}})|)}$, $\mathbf{u}^k = ((u_{1,a}^k, \ldots, u_{H,a}^k), a \in \mathcal{A}^{\mathrm{t}})^\top$, $\widetilde{\boldsymbol{\phi}}^k = ((\widetilde{\phi}_{1,a}^k, \ldots, \widetilde{\phi}_{H,a}^k), a \in \mathcal{A}^{\mathrm{t}})^\top$, and $\boldsymbol{\omega}^k = ((\omega_{1,a}^k, \ldots, \omega_{H,a}^k), a \in \mathcal{A}^{\mathrm{t}})^\top$. We obtain the approximate formulation of the relocation problem ($\mathcal{P}$) as follows:

$$\Theta'(\mathbf{x}) = \min_{\widetilde{\mathbf{Y}}^k,\ k \in \mathcal{K}} \left\{ (7) \;\middle|\; (1b), (1c), (2a)-(2h), (3), \left((5c)-(5e), (6),\ a \in \mathcal{A}^{\mathrm{t}}\right),\ k \in \mathcal{K} \right\}. \tag{$\mathcal{Q}_0$}$$

By relaxing $u_{h,a}^k$ to be continuous for $k \in \mathcal{K}$, $h \in \mathcal{H}$, and $a \in \mathcal{A}^{\mathrm{t}}$, the above problem becomes

$$\begin{aligned}\Theta''(\mathbf{x}) = \min_{\widetilde{\mathbf{Y}}^k,\ k \in \mathcal{K}} \Bigg\{ (7) \;\Bigg|\; & (1b), (1c), (2a)-(2h), (3), \\ & \left((5c)-(5d), (6), \left(u_{h,a}^k \in [0,1],\ h \in \mathcal{H}\right),\ a \in \mathcal{A}^{\mathrm{t}}\right),\ k \in \mathcal{K} \Bigg\}.\end{aligned} \tag{$\mathcal{Q}$}$$

PROPOSITION 2. *Given any initial vehicle allocation $\mathbf{x} \in \mathcal{X}$, we have $\Theta'(\mathbf{x}) = \Theta''(\mathbf{x})$.*

Proposition 2 shows that for each scenario $k \in \mathcal{K}$, the optimal objective value of problem ($\mathcal{Q}_0$) is not affected if we relax $u_{h,a}^k$ $(h \in \mathcal{H}, a \in \mathcal{A}^{\mathrm{t}})$ to continuous variables. Thus, we can use problem ($\mathcal{Q}$) to approximate the second-stage problem ($\mathcal{P}$) to reduce the computational burden. We denote the two-stage problem ($\mathcal{M}$) under the piecewise-linear approximation as problem ($\mathcal{M}_H$). We provide the abstract forms of problem ($\mathcal{M}$) and its resulting approximation ($\mathcal{M}_H$) (with an objective value Γ^H) in Appendix E.2. The following proposition shows that when the number of pieces H in problem ($\mathcal{M}_H$) goes to infinity, Γ^H converges to the original objective value Γ.

PROPOSITION 3. *If there exists $\hat{H} > 0$ such that problems ($\mathcal{M}$) and ($\mathcal{M}_H$) share the same optimal integer solution $(\mathbf{x}^*, \tilde{\mathbf{z}}^*)$ for any $H > \hat{H}$, and problem ($\mathcal{M}$) has a feasible solution $(\cdot)^\dagger$ such that $\mathbf{x}^\dagger = \mathbf{x}^*$, $\tilde{\mathbf{z}}^\dagger = \tilde{\mathbf{z}}^*$, and $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k\dagger}) < B_c$ for any $k \in \mathcal{K}$, then $\lim_{H \to \infty} \Gamma^H = \Gamma$.*

We also present an alternative asymptotic convergence result in Appendix E.4. Hereafter, we replace problem ($\mathcal{P}$) with the reformulated relocation problem ($\mathcal{Q}$) in the integrated allocation and relocation problem ($\mathcal{M}$).

5.2. Solution Algorithm

We propose an algorithm that can efficiently solve large-scale instances of problem ($\mathcal{M}$). The algorithm consists of three steps: First, we reduce the number of binary variables in the relocation problem ($\mathcal{Q}$). Second, as the entire time-space network $\mathcal{G}$ is too large, we develop a temporal decomposition approach by iteratively solving a series of subproblems. Each subproblem considers only a part of the time-space network covering several consecutive periods. Third, based on initial vehicle allocations obtained in the second step, we design a heuristic to further improve the solution quality. To facilitate the algorithm description, we first refine our notation.

Algorithm Notation. Recall from Section 3.2 that the starting periods of trip arcs and 3PL relocation arcs are in the sets $\mathcal{T}(l_{ij})$ and $\mathcal{T}(l_\mathrm{r})$, where l_{ij} and l_r represent the trip and relocation durations, respectively. For recording purposes, we rename the sets as $\mathcal{T}(l_{ij}, 0, T) = \{0, 1, \ldots, T - l_{ij} - 1\}$ and $\mathcal{T}(l_\mathrm{r}, 0, T) = \{0, 1, \ldots, T - l_\mathrm{r} - 1\}$. Likewise, we rename the relocation problem ($\mathcal{Q}$) as $\mathcal{Q}(0, T)$ and the integrated allocation and relocation problem ($\mathcal{M}$) as $\mathcal{M}(0, T)$. Generally, given a starting period $s \in \mathcal{T}$, an ending period $e \in \mathcal{T} \cup \{T\}$ ($e \geq s + 1$), and a time duration $l \in \{1, 2, \ldots, e - s\}$, define $\mathcal{T}(l, s, e) = \{s, s + 1, \ldots, e - l - 1\}$. For a part of the time-space network with a starting period s and an ending period e, let $\mathcal{Q}(s, e)$ denote the corresponding relocation problem with variables $\widetilde{\mathbf{Y}}^k_{(s,e)}$ and a feasible region $\mathcal{Y}_{(s,e)}(\mathbf{x}, \boldsymbol{\lambda}^k) := \{(26a) - (26q)\}$ for $k \in \mathcal{K}$ (see Appendix E.5 for details). Let $\mathcal{M}(s, e)$ denote the integrated allocation and relocation problem with $\mathcal{Q}(s, e)$ as the relocation problem in its second stage.

When we relax the binary variables $\tilde{\mathbf{z}}^k, k \in \mathcal{K}$, to continuous ones, we denote the resulting feasible region in the second stage as $\mathcal{Y}_{\mathrm{LP}(s,e)}(\mathbf{x}, \boldsymbol{\lambda}^k) := \big\{(26a) - (26p),\ \tilde{z}^k_t \in [0, 1],\ t \in \mathcal{T}(l_\mathrm{r}, s, e)\big\}, k \in \mathcal{K}$. The relaxed relocation problem is $\mathcal{Q}_{\mathrm{LP}}(s, e) := \{\Theta''(\mathbf{x}) |\ \widetilde{\mathbf{Y}}^k_{(s,e)} \in \mathcal{Y}_{\mathrm{LP}(s,e)}(\mathbf{x}, \boldsymbol{\lambda}^k),\ k \in \mathcal{K}\}$ and the corresponding two-stage problem is denoted by $\mathcal{M}_{\mathrm{LP}}(s, e)$. Note that $\mathcal{Q}_{\mathrm{LP}}(s, e)$ is a linear programming (LP) relaxation of $\mathcal{Q}(s, e)$. For any starting period s, let $\mathbf{x}^s = (x^s_i,\ i \in \mathcal{V})^\top$ denote the first-stage solution and v^s denote the corresponding objective value of $\mathcal{M}_{\mathrm{LP}}(s, e)$. Similarly, we add a superscript s to the second-stage decision variables to denote the second-stage solution.

Step 1: Reducing The Number of Binary Variables. Due to (2a)–(2e), in an optimal solution, only a limited number of $\tilde{z}^k_t,\ k \in \mathcal{K}, t \in \mathcal{T}$, will be 1, and the others equal 0. This means that the 3PL relocation service is provided only in several periods. To reduce the number of binary variables, we design Algorithm 1 (see Appendix E.6) with some $\tilde{z}^k_t$ fixed at 0, while the others are optimized. Specifically, for any $s \in \mathcal{T}$ and $e \in \mathcal{T} \cup \{T\}$ such that $e - s \geq 3$, we solve $\mathcal{M}_{\mathrm{LP}}(s, e)$ to find $\tilde{z}^{ks}_t$ for each scenario $k \in \mathcal{K}$. If the values of $\tilde{z}^{ks}_t$ and $\tilde{z}^{ks}_t + \tilde{z}^{ks}_{t+1}$ are almost zero, then we fix $\tilde{z}^k_t = 0$ in $\mathcal{Q}(s, e)$.

Step 2: Temporal Decomposition. The original relocation problem $\mathcal{Q}(0, T)$ may have a huge number of variables and constraints. By splitting the time-space network into M sub-networks along the operational horizon, each sub-network $m \in \{1, 2, \ldots, M\}$ has a shorter time horizon with $T_{\text{sub}} = \lfloor T/M \rfloor$ periods. Solving problem $\mathcal{Q}(s,e)$ with $s = (m-1)T_{\text{sub}}$ and $e = mT_{\text{sub}}$, which corresponds to the mth sub-network, is more computationally efficient. Thus, we will solve a sequence of smaller two-stage stochastic programs, each corresponding to a sub-network, in a backward manner from sub-network M to sub-network 1.

For each sub-network m, we follow a four-step solution procedure. First, we apply Algorithm 1 to reduce the number of binary variables. Second, based on the initial vehicle allocation of sub-network $m+1$, we add new constraints to balance the vehicle flows between sub-networks m and $m+1$, and obtain the initial vehicle allocation of sub-network m. Third, we perturb the initial vehicle allocation of sub-network m and re-solve $\mathcal{M}_{\text{LP}}(s,e)$ several times to yield several candidate allocations. Fourth, if $m \geq 2$, then we choose the vehicle allocation with the best objective of $\mathcal{M}_{\text{LP}}(s,e)$ as the vehicle allocation of sub-network m. If $m = 1$, then given an initial vehicle allocation of sub-network m, we solve the two-stage problem $\mathcal{M}(0,T)$ over the original time-space network with the first-stage decisions fixed at this initial vehicle allocation and the number of binary variables reduced. We choose the vehicle allocation with the best objective of $\mathcal{M}(0,T)$ as the initial vehicle allocation of the entire network. Algorithm 2 in Appendix E.7 presents the above decomposition approach in detail to solve problem ($\mathcal{M}$).

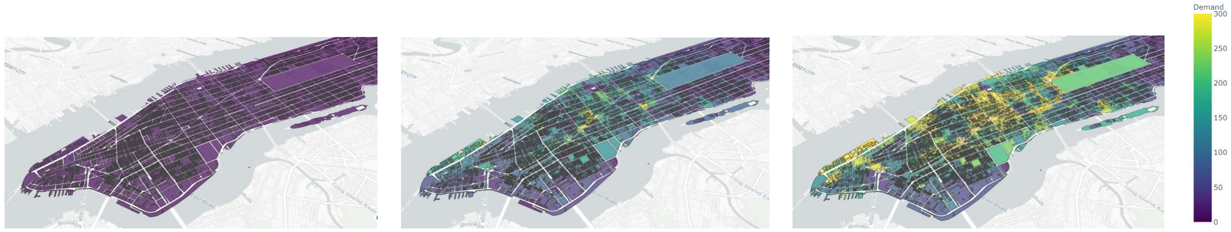
Step 3: Heuristic Search. With several initial vehicle allocations of sub-network 1 obtained from the above temporal decomposition approach, we introduce Algorithm 3 in Appendix E.8 to further search for better solutions. Algorithm 3 is based on the following intuition: If the objective value under a vehicle allocation monotonically decreases as the total number of allocated vehicles increases (resp. decreases), then likely a smaller objective can be found by further increasing (resp. decreasing) the total number of allocated vehicles. We observe this phenomenon in Section 6.

6. Numerical Experiments: A Case Study

We conduct numerical experiments using real data from Citi Bike (2021). We first discuss parameter settings and then obtain managerial insights from various experiments based on these settings.

6.1. Parameter Settings

We have collected data from Citi Bike (2021) in NYC from January 1, 2018 to December 31, 2019. Figure 5 illustrates the daily demands of different locations of Manhattan in 2018. We focus on Midtown Manhattan between the 20th and 57th Streets, and divide the area into $|\mathcal{V}| = 9$ service regions. Each region covers a rectangular area of about 1 km^2 as shown in Figure 6.[1] The travel distance between any two regions is measured by the Manhattan distance (L1 norm) between their centers. Based on the data in 2018, the average riding speed is 9 km/hour.[2] The average trip duration from one region to a neighboring region is 6.7 minutes. We assume the traveling speed between any two regions is constant and set each period as 6 minutes (leading to $T = 240$ periods per day). The longest trip (e.g. from region 1 to region 9) takes four periods. This is evidenced in the data as 96.5% of the trips in the studied area finish within 24 minutes.



(a) Minimum Daily Demands **(b)** Average Daily Demands **(c)** Maximum Daily Demands

Figure 5 Daily Demands of Manhattan in 2018

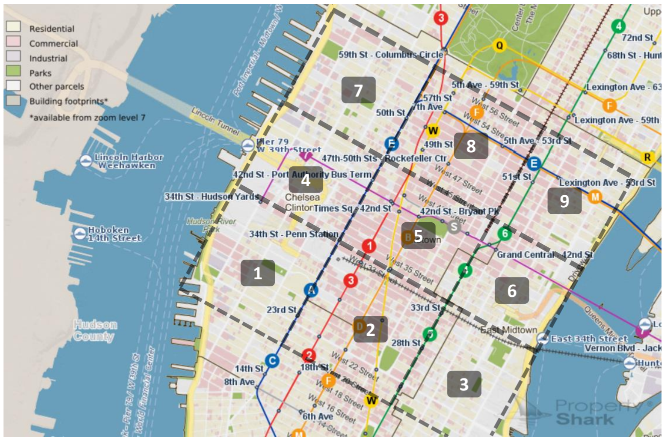


Figure 6 Service Regions in Midtown Manhattan

The upper bound N of the total number of vehicles allocated to the studied area is not available in the data set. We estimate such an upper bound as $N = \frac{\text{number of all the trips in the studied area}}{\text{number of all the trips in NYC}} \times$ number of vehicles in NYC, where the number of vehicles in NYC is obtained by counting unique bike IDs in the data set. Based on the above calculation, we have $N = 1{,}206$. Similarly, we estimate the upper bound of the number of vehicles allocated to region j as $B_j = \frac{\text{number of all the trips from region } j}{\text{number of all the trips in the studied area}} \times N,\ j \in \mathcal{V}$. Note that the total number of vehicles allocated by our model is much smaller than N, indicating that many redundant bikes are allocated to the studied area — a phenomenon commonly seen in practice. Having redundant bikes creates traffic congestion on the streets and leads to more complicated vehicle relocation operations. Thus, we reduce N from 1,206 to 500. The upper bound B_j, $j \in \mathcal{V}$, is also scaled down by multiplying $500/1{,}206$.

The cost parameters are estimated in USD. The vehicle allocation cost is $c_j = 0.5$ for $j \in \mathcal{V}$.[3] The revenue per bike trip per period is $R = 0.2$ and the penalty cost $C_{\mathrm{p}} = 0.5$ for an unfulfilled trip demand.[4] We set $C_{\mathrm{r}} = 15$ for each request of 3PL relocation with volume lower capacity $\underline{q} = 0$ and upper capacity $\bar{q} = 100$.[5] The 3PL can be requested at most $\bar{z} = 10$ times per day. In any given $l_{\mathrm{f}} = 10$ consecutive periods, the 3PL can operate at most $\bar{\bar{z}} = 2$ periods.

We now obtain the parameters of rider crowdsourcing. Given any regions $i, j \in \mathcal{V}$, we construct the piecewise-linear approximation of the incentive function in (5a)–(5e) using three parameters $\bar{\Lambda}_{ij}$, $\bar{k}_{hij}$, and $\bar{\phi}_{ij}$. We estimate these parameters as follows. (i) We first solve the two-stage problem ($\mathcal{M}$) over the entire network $\mathcal{G}$ without any relocations. Denote its solution with superscript "†", and set $\bar{\Lambda}_{ij} = \max\{\eta^\dagger_{n_{it}, n_{j,t+l_{ij}}} \mid t \in \mathcal{T}(l_{ij})\}$. (ii) We generate H slopes $\bar{k}_{hij}$, $h \in \{1, 2, \ldots, H\}$, of the piecewise-linear function by obtaining the corresponding unit reward $1/\bar{k}_{hij}$. The H unit rewards are sampled from a uniform distribution $U(\underline{u}_{ij}, \bar{u}_{ij})$, where $\underline{u}_{ij} = 0.1 + 0.1 l_{ij}$ and $\bar{u}_{ij} = 0.2 + 0.1 l_{ij}$ are estimated using the data of the Citi Bike's Bike Angels Rewards Program.[6] We then sort the H slopes in decreasing order to fit the incentive function in Figure 4. (iii) We derive the length of each segment in Figure 4 as $\delta_{ij} = (\bar{\Lambda}_{ij} - \epsilon) / \sum_{h=1}^{H} \bar{k}_{hij}$ and the reward threshold as $\bar{\phi}_{ij} = H\delta_{ij}$. The trip duration between any two regions $i, j \in \mathcal{V}$ ranges from 1 to 4 periods. The maximum reward for crowdsourcing one rider is 0.6 dollar (i.e., $\max\{\bar{u}\} = \max\{0.2 + 0.1 l_{ij} \mid l_{ij} \in \{1, \ldots, 4\}\} = 0.6$). We set the crowdsourcing budget $B_c = 500$ as the actual crowdsourcing costs in our experiments are always lower than this value.

The data reveals that the hourly total demand of the studied area exhibits different patterns on weekdays and weekends. The former has a bimodal pattern with morning and evening peak hours, while the latter has a unimodal pattern with afternoon peak hours. Thus, we conduct the case study for the weekdays and weekends separately. In the experiments, we use the 2018 data for training to compute the first-stage allocation $\mathbf{x}$ by solving problem ($\mathcal{M}$). Given $\mathbf{x}$, we use the 2019 data to perform out-of-sample tests by solving problem ($\mathcal{Q}$). We use 350 days of trip records each year as 350 different scenarios (i.e., one day corresponds to one scenario) to capture the demand uncertainty, where 250 scenarios correspond to the weekday demands and 100 scenarios correspond to the weekend demands. We set the number of sub-networks $M = 8$. An efficient system should have a low total demand loss $\sum_{a \in \mathcal{A}^{\text{t}}} \eta_a$ and a high expected vehicle utilization rate $\sum_{a \in \mathcal{A}^{\text{t}}} (y_a - \Lambda_a) / \sum_{i \in \mathcal{V}} x_i$. We use these two measures to evaluate the service quality.

6.2. Computational Performance of Our Solution Approach

We demonstrate the computational efficiency of our solution approach in Section 5.2 for solving problem ($\mathcal{M}$) by benchmarking it against CPLEX 12.71 with C++ API.[7] We implement our solution approach in parallel on 24 threads with C++ OpenMP and perform computational experiments on a computing node with 24 2.3-GHz Intel Xeon E5-2670 processors and 32 GB of memory in a high performance computing cluster.

Note that more scenarios lead to a higher solution quality but require more computational time. Figure 7 shows how the relative error of the total allocation[8] and the computational time vary with the number of scenarios $|\mathcal{K}|$ under our solution algorithm. Both approaches above cannot handle 250 scenarios for weekday training instances because of out of memory. However, as $|\mathcal{K}|$ decreases, the computational time decreases whereas the relative error first decreases and then increases. To strike a balance between the computational time and the relative error, we randomly select 80 out of the 250 scenarios and observe that the resulting relative error is only 1.18% on average. Note that the benchmark CPLEX approach cannot even solve any weekday training instance with 80 scenarios.

We do not reduce the number of scenarios for the weekend training instances because our solution approach can efficiently solve all of them. In contrast, the benchmark CPLEX approach struggles to get a high-quality solution. We set 180 minutes as the computational time limit for all the instances in our experiments if not specified.

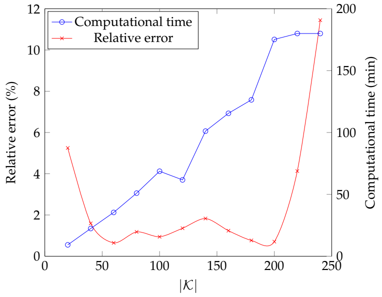


Figure 7 The Trade-off between the Computational Time and the Relative Error

We demonstrate the computational efficacy of our solution approach by varying several parameters. We consider (i) the vehicle allocation cost $c_j \in \{0.5, 0.8, 1.0\}$ for $j \in \mathcal{V}$;[9] (ii) the logistics market fluctuation such that the 3PL relocation cost is $C_{\rm r} \in \{13, 15, 17\}$;[10] (iii) different market competition levels such that the penalty cost $C_{\rm p} \in \{0.1, 0.3, 0.5, 0.7\}$. This results in 36 $(= 3 \times 3 \times 4)$ problem instances for a given weekday or weekend demand setting. For each instance, we run three random experiments with different demand scenarios and report the average result in Table F2 in Appendix F.1.

Our solution approach significantly outperforms the benchmark CPLEX approach for solving problem $(\mathcal{M})$ in terms of the profit and computational time for all the instances (see Table F2). The CPLEX approach cannot provide any feasible solution within the time limit of 180 minutes for the weekday instances, and can only provide a solution $x_j = 0,\ j \in \mathcal{V}$, for the weekend instances. In contrast, our solution approach produces high-quality solutions within 37.7 minutes on average for each weekday instance, and within 45.9 minutes on average for each weekend instance. Furthermore, we find that the optimality gap between the profits of our solution approach and the original problem $(\mathcal{M})$ is around 4% (see Tables F3–F5). These results suggest that our solution approach is promising to support urban shared micromobility operations. We investigate its out-of-sample performance and obtain managerial insights in the following sections.

6.3. Impact of Initial Vehicle Allocation

We examine the impact of the initial vehicle allocation on the system performance under the parameter settings in Section 6.1. We first vary the allocation upper bound (UB) N between 120 and 720, while keeping other parameters unchanged. We focus on the weekday demands as similar insights are obtained from the weekends. Specifically, we examine how N affects the initial vehicle allocation in Figure 8 and the average number of relocated vehicles in Figure 9. Figure 8 suggests that when N is large (i.e., $N > 210$), it is economical to allocate fewer vehicles to each region than its upper bound. There is no need to allocate too many vehicles if demands can be well satisfied. It is interesting to see that more vehicles are allocated to regions 1 to 5, where public transportation and commercial blocks are denser than other regions (see Figure 6).[11]

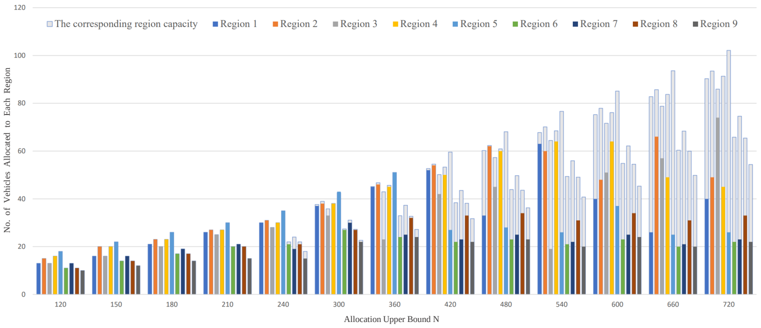


Figure 8 The Impact of Allocation UB on Vehicle Allocation for Different Regions

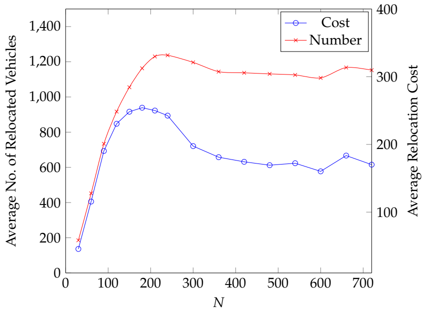


Figure 9 The Impact of Allocation UB on Vehicle Relocation

Figure 9 suggests that as N increases, both the average number of relocated vehicles and the average relocation cost first increase and then decrease. When the UB is low (i.e., $N \leq 240$), the initially allocated vehicles are not enough to satisfy the demands. Raising the UB increases the initial allocation (see Figure 8) and leads to more relocated vehicles. Thus, the number of relocated vehicles increases as the initial allocation increases. When the allocation UB is moderate (i.e., $240 < N \leq 480$), a larger N decreases both the average number of relocated vehicles and the average relocation cost. This is because a larger N leads to a higher initial allocation (see Figure 8) and the allocated vehicles are sufficient to satisfy most of the demands, which requires less vehicle relocation subsequently. This suggests that a higher initial allocation can reduce the need for subsequent relocation. When the allocation UB is large (i.e., $N > 480$), both the initial allocation and subsequent relocation stabilize and are less affected by the UB. Therefore, we choose a sufficiently large $N = 500$ for the training instances. The optimal initial vehicle allocation $\mathbf{x}^*$ results in 338 vehicles in total for the weekdays and 253 for the weekends. Such initial vehicle allocations will be used for later experiments.

We further examine the impact of the initial allocation by multiplying $\mathbf{x}^*$ with $1+\alpha$, where α represents a variation ratio. For each α, we solve the relocation problem $(\mathcal{Q})$ through an out-of-sample test. The profit gap in the last column of Table 1 is defined as $\frac{(\text{profit with }_{\alpha=0})-(\text{profit with }_{\alpha\neq 0})}{\text{profit with }_{\alpha=0}} \times 100\%$. Table 1 shows that it is crucial to optimize the initial allocation as it has great impact on the system performance. The profit drops when the initial allocation deviates from the optimal allocation $\mathbf{x}^*$. Specifically, if the initial allocation decreases by 75%, the profit drops by 52.20% on

Table 1 Impact of Initial Vehicle Allocation

	Variation Ratio α	Crowdsourcing		3PL		Demand Loss	Utilization Rate	Profit ($)	Profit Gap (%)
		No. of Vehicles Relocated	Cost ($)	No. of Vehicles Relocated	Cost ($)				
Weekdays	-0.75	689.06	178.17	17.14	3.78	1,448.56	65.32	1,186.72	**52.20**
	-0.50	877.40	222.64	256.14	36.00	295.64	39.21	2,204.52	11.21
	-0.25	732.79	181.58	518.62	55.86	31.02	27.44	2,432.84	2.01
	0	491.50	114.80	673.20	57.30	0	20.71	**2,482.73**	-
	0.25	353.75	80.19	674.99	51.00	0	16.55	2,481.09	0.07
	0.50	248.39	54.45	669.24	47.28	0	13.75	2,467.55	0.61
	0.75	186.58	40.20	618.89	42.18	0	11.83	2,445.40	1.50
Weekends	-0.75	588.94	157.93	0.63	0.15	249.13	48.08	961.86	**18.31**
	-0.50	479.76	129.82	155.95	26.70	0.76	26.15	1,168.89	0.73
	-0.25	336.59	85.44	249.97	30.30	0	17.76	1,177.13	0.03
	0	246.90	58.80	302.30	27.80	0	13.34	**1,177.50**	-
	0.25	184.20	43.49	306.74	23.40	0	10.61	1,164.70	1.09
	0.50	135.27	31.43	311.10	21.60	0	8.83	1,146.56	2.63
	0.75	105.09	24.30	296.83	21.00	0	7.62	1,123.79	4.56

weekdays and by 18.31% on weekends because of insufficient vehicles. Furthermore, as the initial allocation varies, the relocation by crowdsourcing and the 3PL also changes. For instance, as α increases on weekdays, the vehicle relocation by crowdsourcing first increases then decreases. This is because when the initial number of vehicles is small, not many vehicles are available for relocation. As the initial number of vehicles increases, more vehicles are available for relocation by crowdsourcing. However, as the initial allocation further increases, it becomes justifiable for mass relocation, which raises the vehicle relocation by the 3PL and reduces that by crowdsourcing. As the initial allocation further increases, sufficient vehicles are available in the system to serve the demands, reducing the needs for relocation by both crowdsourcing and the 3PL.

6.4. Impact of Relocation Strategies

We examine four different relocation strategies for problem ($\mathcal{Q}$): (i) No vehicle relocation, i.e., $\phi_{n_{it},n_{j,t+l_{ij}}} = \Lambda_{n_{it},n_{j,t+l_{ij}}} = 0$ for $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$, and $z_t = 0$ for $t \in \mathcal{T}(l_{\mathrm{r}})$; (ii) relocation only by rider crowdsourcing, i.e., $z_t = 0$ for $t \in \mathcal{T}(l_{\mathrm{r}})$; (iii) relocation only by the 3PL, i.e., $\phi_{n_{it},n_{j,t+l_{ij}}} = \Lambda_{n_{it},n_{j,t+l_{ij}}} = 0$ for $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$; and (vi) relocation by both rider crowdsourcing and the 3PL. We conduct experiments over various instances to assess their performance. Table 2 shows that demand loss can be effectively reduced by vehicle relocation. Compared to the case without any relocation, the demand loss can be significantly reduced by 98% through crowdsourcing relocation alone and 50% through 3PL relocation alone, and can be further reduced when both relocation methods are used. Furthermore, relocation only by crowdsourcing yields a higher utilization rate and profit compared to relocation only by the 3PL. A more detailed observation reveals that crowdsourcing alone relocates more vehicles than the 3PL alone. *Thus, compared to the 3PL alone, rider crowdsourcing alone is more effective to match supply with demand for the shared micromobility system.*

Table 2 Impact of Relocation Strategies and Penalty Cost

	C_p	Relocation Strategies	Crowdsourcing		3PL		Demand Loss	Utilization Rate	Profit ($)
			No. of Vehicles Relocated	Cost ($)	No. of Vehicles Relocated	Cost ($)			
Weekdays	0.1	No relocation	–	–	–	–	953.0	17.892	2,209.9
		Crowdsourcing	928.8	238.6	–	–	24.2	20.640	2,407.8
		3PL	–	–	650.7	63.5	473.3	19.311	2,398.2
		Combination	484.0	111.7	660.1	57.0	14.8	20.668	**2,481.4**
	0.3	No relocation	–	–	–	–	947.2	17.909	2,019.9
		Crowdsourcing	960.6	247.8	–	–	1.9	20.706	2,406.1
		3PL	–	–	665.1	66.7	457.0	19.359	2,309.5
		Combination	492.1	115.2	673.2	56.9	0.3	20.711	**2,482.5**
	0.5	No relocation	–	–	–	–	942.8	17.922	1,830.9
		Crowdsourcing	964.2	248.7	–	–	0.1	20.711	2,405.9
		3PL	–	–	673.3	67.6	455.3	19.364	2,217.1
		Combination	491.5	114.8	673.2	57.3	0	20.711	**2,482.7**
	0.7	No relocation	–	–	–	–	942.7	17.922	1,642.3
		Crowdsourcing	964.2	248.8	–	–	0.1	20.711	2,405.9
		3PL	–	–	677.6	69.1	451.1	19.377	2,129.9
		Combination	489.4	113.9	674.1	57.6	0	20.711	**2,483.3**
Weekends	0.1	No relocation	–	–	–	–	494.5	11.383	1,008.7
		Crowdsourcing	492.0	133.9	–	–	0.4	13.336	1,130.0
		3PL	–	–	318.5	33.8	218.1	12.475	1,128.4
		Combination	228.5	53.9	320.2	30.2	0.8	13.334	**1,179.8**
	0.3	No relocation	–	–	–	–	492.7	11.390	910.0
		Crowdsourcing	492.4	134.1	–	–	0.1	13.337	1,130.0
		3PL	–	–	319.8	34.5	219.2	12.471	1,082.8
		Combination	235.3	56.3	305.6	29.1	0	13.337	**1,178.6**
	0.5	No relocation	–	–	–	–	491.2	11.396	811.6
		Crowdsourcing	492.5	134.1	–	–	0	13.337	1,130.0
		3PL	–	–	309.3	34.8	228.5	12.434	1,030.4
		Combination	246.9	58.8	302.3	27.8	0	13.337	**1,177.5**
	0.7	No relocation	–	–	–	–	491.2	11.396	713.4
		Crowdsourcing	492.5	134.1	–	–	0	13.337	1,130.0
		3PL	–	–	301.1	34.5	227.1	12.440	985.0
		Combination	245.6	59.1	302.9	27.9	0	13.337	**1,177.0**

To understand the above results, it is helpful to distinguish two types of relocation needs: *mass relocation needs* (i.e., many vehicles to be relocated, such as in rush hours) and *sporadic relocation needs* (i.e., only a few vehicles to be relocated, such as in non-rush hours). Note that rider crowdsourcing serves both the mass and sporadic relocation needs, while the 3PL generally only serves the mass relocation needs. The 3PL is not cost effective for the sporadic relocation because its average cost per relocated vehicle is much higher than the cost of losing a customer. The 3PL achieves a lower average cost per relocated vehicle for the mass relocation because of the economies of scale in relocating vehicles in batches. Furthermore, Table 2 shows that when both relocation methods are used, the 3PL relocates more vehicles with a lower cost than rider crowdsourcing. The number of vehicles relocated by the 3PL is similar to that when only 3PL relocation is used. This suggests that when both methods are employed, the 3PL satisfies mostly the mass relocation needs, while rider crowdsourcing mainly serves the sporadic relocation needs.

In general, as the penalty cost per customer lost C_p increases, it becomes more likely to relocate vehicles to avoid demand losses. However, Table 2 shows an opposite effect for relocation only by the 3PL on weekends: Raising C_p may decrease the relocation by the 3PL.[12] Figure 10 shows a typical scenario on weekends: If the penalty cost is low ($C_p = 0.1$), the 3PL relocates some vehicles (47 in this example) around 5pm. Surprisingly, if the penalty cost is high ($C_p = 0.5$), it does not relocate any vehicles. This counter-intuitive situation only happens on weekends when the demands and the need for relocation are lower. The unimodal pattern of customer arrivals over a day in Figure 10 shows fewer rush hours, reducing the need for mass relocation. For example, if $C_p = 0.5$, the 3PL is requested 4.5 times on a weekday ($\approx$ total relocation cost by the 3PL$/C_r = 67.6/15$) and is requested only 2.3 times on a weekend ($\approx 34.8/15$). As C_p increases, the profit gain from region j may be less than the penalty cost incurred in region i, reducing the relocation by the 3PL from region i to region j. Furthermore, Table 2 shows that as C_p increases on weekends, the profit gap between 3PL relocation alone and the other two strategies ("Crowdsourcing" and "Combination") increases. Therefore, the operator should use either rider crowdsourcing alone or the combination of both methods instead of using the 3PL alone to relocate vehicles on weekends.

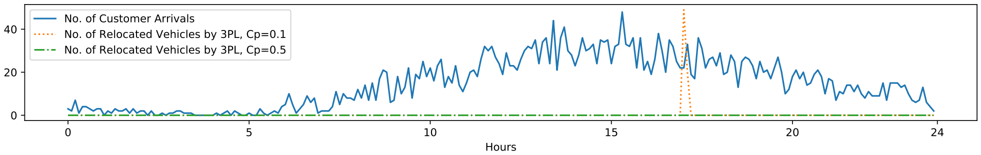


Figure 10 Customer Arrivals and Vehicles Relocated by The 3PL in A Day during Weekends

6.5. Impact of Rider Crowdsourcing Budget

Table 3 Impact of Rider Crowdsourcing Budget

	B_c	Crowdsourcing		3PL		Demand Loss	Utilization Rate	Profit ($)
		No. of Vehicles Relocated	Cost ($)	No. of Vehicles Relocated	Cost ($)			
Weekdays	0	0	0	673.3	67.6	455.3	19.365	2,217.1
	30	140.0	29.0	693.4	67.0	315.7	19.777	2,310.5
	50	223.4	46.9	689.5	66.2	234.4	20.018	2,357.9
	80	323.5	69.8	694.2	64.4	140.6	20.296	2,409.3
	100	373.7	82.1	685.8	62.9	95.2	20.430	2,433.4
	∞	491.5	114.8	673.2	57.3	0	20.711	2,482.7
Weekends	0	0	0	309.3	34.8	228.5	12.435	1,030.4
	10	47.7	9.7	316.0	34.2	181.4	12.621	1,064.6
	20	89.4	18.4	312.4	33.8	140.9	12.780	1,091.6
	30	125.4	26.4	313.3	32.4	108.2	12.910	1,112.5
	50	176.6	38.7	307.3	30.8	61.8	13.093	1,140.7
	∞	246.9	58.8	302.3	27.8	0	13.337	1,177.5

The shared micromobility operator often contracts with the 3PL for its relocation service with a stable relocation budget. In contrast, the operator pays a crowdsourced rider per trip and has a more flexible budget for rider crowdsourcing. The operator can experiment and adjust the crowdsourcing budget B_c to learn its effect. Table 3 shows the impact of B_c on the system performance. As B_c increases, more riders are crowdsourced and the relocation by these riders significantly increases. Meanwhile, the demand loss decreases and the vehicle utilization rate increases, leading to a higher profit. In contrast, as B_c increases, the relocation by the 3PL first increases and then decreases in general. The above results can be summarized as follows: *If the crowdsourcing budget is low, rider crowdsourcing complements (benefits) the 3PL. On the other hand, if the budget is high, rider crowdsourcing substitutes the 3PL.*

Figure 11 illustrates how rider crowdsourcing can complement the 3PL. The demand from region j in period t to region m is large. If no crowdsourcing is offered in region m due to an insufficient budget, the 3PL may not relocate vehicles from region i in period $t - l_r$ to region j because the demand in region i later (in period $t + l_{jm} + l_{mi}$) is also large. Conversely, if rider crowdsourcing is provided in region m, then the 3PL can relocate vehicles from region i in period $t - l_r$ to region j so that these vehicles can satisfy the large demand from region j in period t to region m. Crowdsourced riders can help relocate these vehicles back to region i to meet its demand in period $t + l_{jm} + l_{mi}$. Thus, rider crowdsourcing can complement the 3PL.

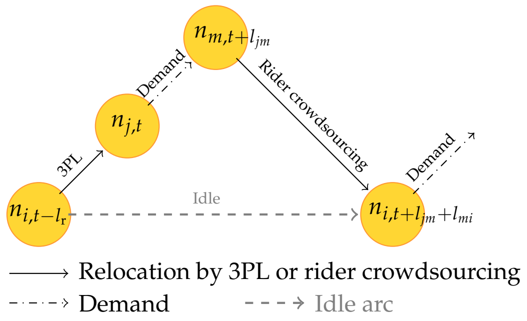


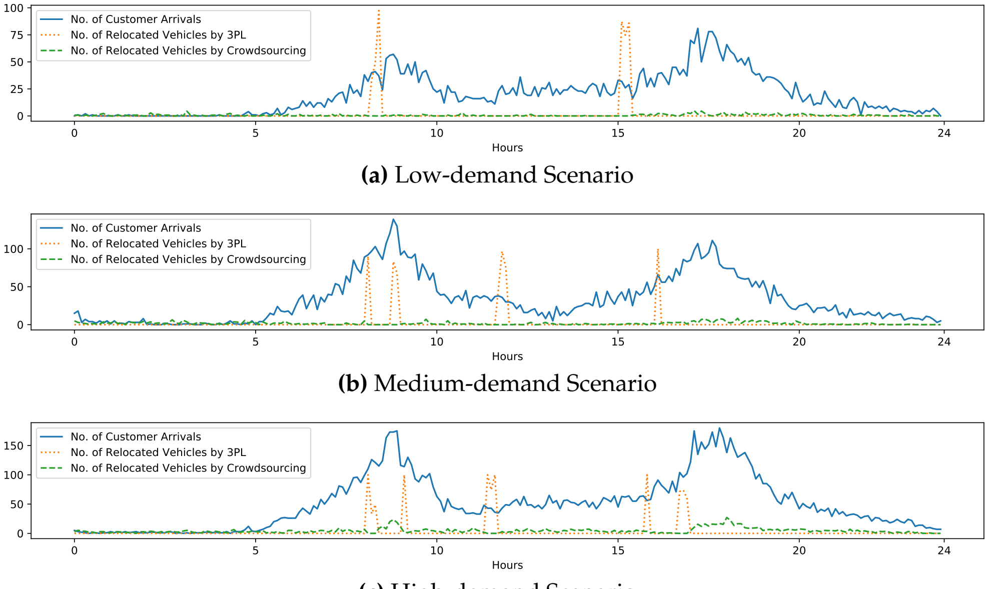
Figure 11 Rider Crowdsourcing Benefits The 3PL

As the crowdsourcing budget becomes larger, the 3PL relocation with a relatively high average cost per relocated vehicle can be substituted by the cheaper rider crowdsourcing, manifesting the substitutional effect. Although the above complementary and substitutional effects can potentially exist, Table 3 shows that as B_c increases, the relocation by the 3PL fluctuates only within 3.1% on weekdays and 4.5% on weekends. This suggests that *introducing rider crowdsourcing to the micromobility system with 3PL relocation may significantly increase the system's profit without much affecting the existing commitment with the 3PL.*

6.6. Operational Features of Relocation

Table 2 shows that the micromobility system generates more profit if both relocation methods are adopted. Here, we examine the operational features of relocation in detail on three representative scenarios for weekdays: the daily demand is low (4,744 trips), medium (7,967 trips), and high (12,294 trips). We observe similar relocation results on weekends.

Temporal Features of Relocation. Figure 12 shows the vehicle relocation compared to the number of customer arrivals over the operational horizon. Clearly, the 3PL often relocates vehicles in batches during 8am-10am, 11am-1pm, and 3pm-5pm (i.e., around peak hours). As the 3PL relocates vehicles to adjust the supply across multiple regions to match potential demands in the upcoming periods, the number of relocated vehicles can be larger than the demand per period. Note that the 3PL is usually not engaged after the evening demand peak because of fewer customer arrivals that do not need batch relocation. In contrast, rider crowdsourcing is conducted throughout the day to relocate a much smaller number of vehicles per period.

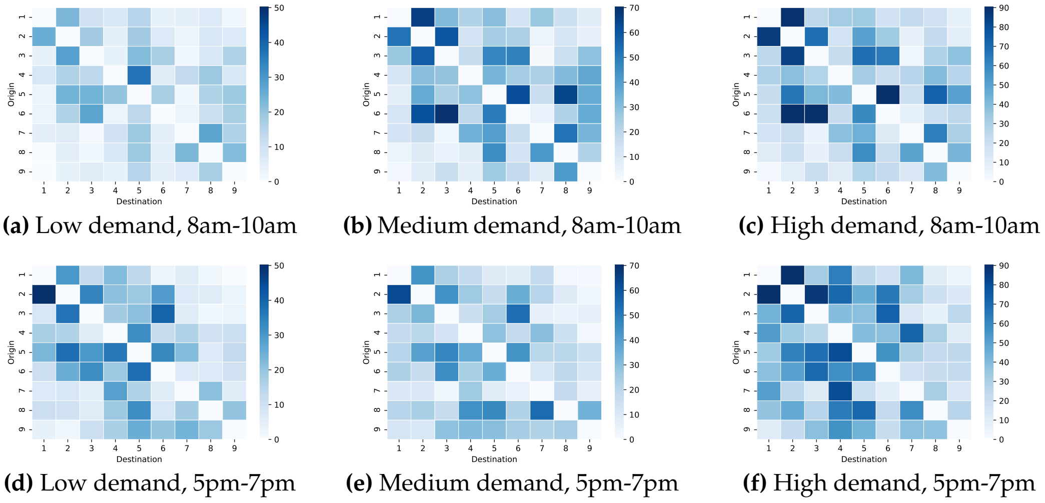


(a) Low-demand Scenario

(b) Medium-demand Scenario

(c) High-demand Scenario

Figure 12 Customer Arrivals and Vehicles Relocated by The 3PL and Rider Crowdsourcing during A Weekday



(a) Low demand, 8am-10am **(b)** Medium demand, 8am-10am **(c)** High demand, 8am-10am

(d) Low demand, 5pm-7pm **(e)** Medium demand, 5pm-7pm **(f)** High demand, 5pm-7pm

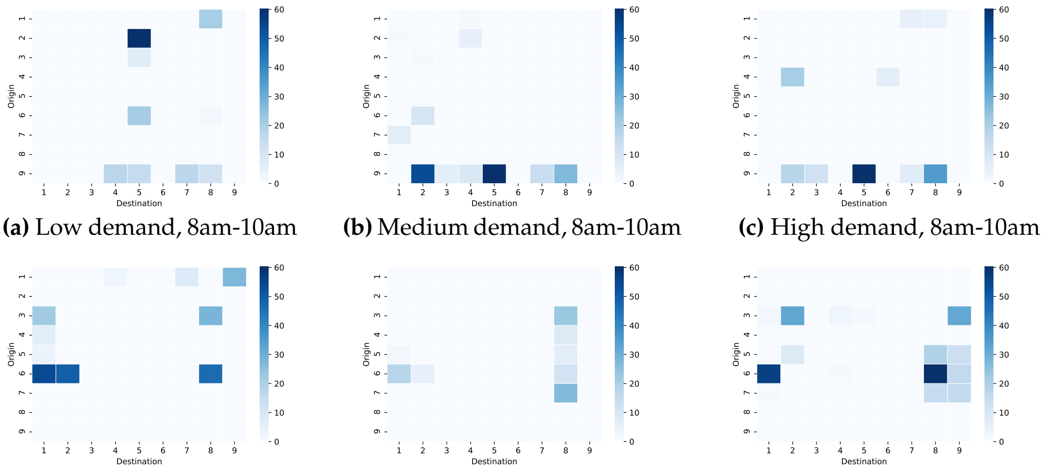
Figure 13 Spatial Features of Demands in Rush Hours

Spatial Features of Relocation. Figure 13 displays the weekday demands during rush hours in the morning (top row) and evening (bottom row). Specifically, regions 2, 4, 5, 6, and 8 represent Midtown South, West, Center, East, and North, respectively, whereas regions 1, 3, 7, and 9 represent Midtown Southwest, Southeast, Northwest, and Northeast, respectively (see Figure 6).

A ridership report of subway traffic in Manhattan (MTA 2020) indicates that region 5 (Midtown Center) is the busiest region with the largest passenger flow. It covers Times Square, Herald

Square, and the north part of Broadway, and has multiple subway lines passing through. Region 6 (Midtown East) has an important transportation hub, the Grand Central Terminal. Region 2 (Midtown South) has high-traffic subway stations, 34St Herald and 34St Penn Station, and is slightly busier than region 8 (Midtown North) where Rockefeller Center is located. Region 1 (Midtown Southwest) covers a few entrances of 34St Penn Station and hence can be a busy region. Region 9 (Midtown Northeast) has one high-traffic station, Lexington Av/53St. Figure 13 suggests that the shared micromobility demand is closely related to the ridership of subway traffic. The demand for short-range trips ($l_{ij} \leq 2$) is higher than that for long-range trips ($l_{ij} \geq 3$).

Figures 14 and 15 show the relocation by the 3PL and rider crowdsourcing, respectively. We focus on the demand features in two clusters of regions and the corresponding relocation features. In the first cluster, the busiest pick-up regions are regions 1, 2, and 3 (respectively, Midtown Southwest, South, and Southeast) during both 8am-10am and 5pm-7pm (see Figure 13), as most customers travel from one of these regions to a neighboring one. Thus, they are also among the busiest destinations. To maintain the vehicle supply, the operator relocates many vehicles to regions 1, 2, and 3 rather than from them. In particular, the 3PL may relocate vehicles from a faraway region (e.g., from region 9 to region 2 in Figure 14(b)). In contrast, rider crowdsourcing relocates from nearby regions (e.g., from region 5 to region 2 in Figure 15(f)).



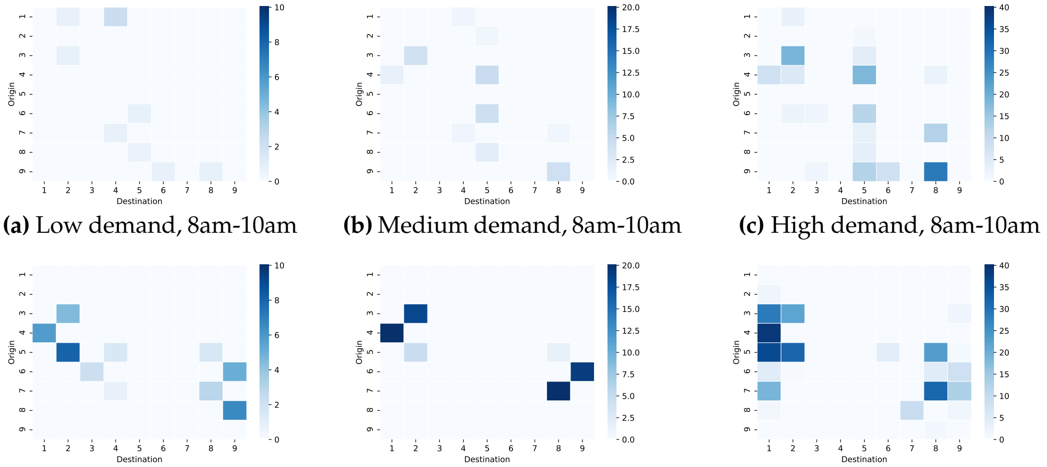
(a) Low demand, 8am-10am **(b)** Medium demand, 8am-10am **(c)** High demand, 8am-10am

(d) Low demand, 3pm-5pm **(e)** Medium demand, 3pm-5pm **(f)** High demand, 3pm-5pm

Figure 14 Spatial Features of 3PL Relocation

In the second cluster, many customers pick up vehicles from region 5 (Midtown Center) and travel to regions 6 and 8 (respectively, Midtown East and North) in the morning (see Figures 13(b) and (c)). Thus, many vehicles are relocated to region 5 from other regions before the customers arrive. In particular, the 3PL relocates batches of vehicles mostly from the low-demand region 9 (Midtown Northeast) to region 5 (see Figures 14(b) and (c)). In contrast, rider crowdsourcing relocates only a few vehicles from each neighboring region to region 5 (see Figures 15(b) and (c)).

These observations suggest that the destinations of relocation by the 3PL and crowdsourcing are similar, indicating their common goal to match supply with demand. The 3PL relocates many vehicles per request from faraway regions, whereas crowdsourcing often incentivizes a few riders each time from neighboring regions. Thus, the two methods are often mutually exclusive at the origins of relocation. This is consistent with Proposition 1, which generally still holds for the case with multiple regions (see Appendix F.2 for details).



(a) Low demand, 8am-10am **(b)** Medium demand, 8am-10am **(c)** High demand, 8am-10am

(d) Low demand, 5pm-7pm **(e)** Medium demand, 5pm-7pm **(f)** High demand, 5pm-7pm

Figure 15 Spatial Features of Rider Crowdsourcing Relocation

6.7. Impact of The Temporal Demand Pattern

Table 4 Impact of The Temporal Demand Pattern

Demand Pattern	C_p	**Crowdsourcing**		**3PL**		**Demand Loss**	**Utilization Rate**	**Profit ($)**
		No. of Vehicles Relocated	**Cost ($)**	**No. of Vehicles Relocated**	**Cost ($)**			
Bimodal (Weekdays)	0.05	254.8	60.4	130.4	17.1	8.4	29.861	1,590.2
	0.1	271.8	65.5	113.5	14.8	3.8	29.888	1,588.4
	0.3	274.1	65.9	113.3	15.0	0.2	29.901	1,588.9
	0.5	264.5	63.3	127.1	17.5	0	29.911	1,589.1
Unimodal (Weekends)	0.05	279.5	64.7	93.7	11.9	9.1	31.879	1,594.9
	0.1	278.6	64.5	101.0	12.8	6.0	31.898	1,594.8
	0.3	291.5	67.9	82.4	10.6	2.8	31.918	1,593.8
	0.5	293.3	68.2	81.9	10.8	2.0	31.923	1,593.4

We now investigate the impact of the temporal demand pattern on vehicle relocation by comparing the bimodal and unimodal customer arrival patterns on weekdays and weekends respectively. Note that weekends have a lower daily demand than weekdays. To see the impact caused only by the demand pattern, we process our data using ARIMA models (Hamilton 2020) so that the daily demand of weekends is on the same scale as the daily demand of weekdays (see Appendix F.3 for details). After training with the new data by solving problem ($\mathcal{M}$), the optimal total number of vehicles allocated under the bimodal pattern is 170, and that under the unimodal pattern is 159. We then solve problem ($\mathcal{Q}$) using testing data to obtain out-of-sample results. Table 4 presents the optimal vehicle relocation under the two demand patterns. We obtain the following insights.

First, the combined number of vehicles relocated by the two methods remains similar under the two demand patterns. However, the 3PL relocates more vehicles under the bimodal pattern than under the unimodal pattern. This is because the two demand peaks of the bimodal pattern require more batch relocations from the 3PL than the unimodal case. In contrast, rider crowdsourcing relocates more vehicles under the unimodal pattern than under the bimodal pattern. Second, the system has more vehicles allocated initially for the bimodal demand pattern (i.e., 170) than that for the unimodal pattern (i.e., 159). However, the profit under the bimodal pattern is lower than that under the unimodal pattern, indicating that the operations under the bimodal demand pattern are more challenging to manage. Since the total costs under the two demand patterns are similar, the higher utilization rate (higher revenue) under the unimodal pattern leads to a higher profit.

7. Conclusion

The operator of a shared micromobility system often bears the cost of physical assets – micromobility vehicles. Thus, it is especially crucial to properly set the initial vehicle allocation for different service regions to satisfy customer demands. In addition, as customers can pick up and drop off vehicles in any region, the system may experience a severe mismatch between vehicle supply and demand under uncertain customer arrivals. To better match supply with demand, the operator needs to constantly relocate the micromobility vehicles across the regions. When the operator plans for the initial vehicle allocation, he needs to take the future relocation decisions into consideration. Thus, it is critical to integrate the initial vehicle allocation and subsequent relocation decisions to maximize the operator's profitability. The problem becomes especially complex when the operator employs both rider crowdsourcing and 3PL for vehicle relocation.

We formulate an integrated vehicle allocation and relocation problem under demand uncertainty as a two-stage stochastic integer program on a time-space network. The operator's objective is to maximize his expected profit. In the first stage, the operator decides the initial vehicle allocation before the demands are realized. In the second stage where demand realizations are observed, the operator decides the relocation of vehicles across the regions to match supply with demand over multiple time periods. For a special case with two service regions, we find that under the optimal relocation strategy, rider crowdsourcing and the 3PL are not used simultaneously in any region and period. Instead, the optimal relocation strategy will choose a method with a lower relocation cost per vehicle (see Proposition 1). For a more general case with multiple service regions, we first approximate the rider incentive function using a piecewise-linear function. The approximation introduces some binary variables, which can be relaxed to continuous variables without affecting the optimal objective value of the second-stage relocation problem (see Proposition 2). We show that the approximation is asymptotically tight (see Proposition 3). We then propose an algorithmic approach that incorporates both scenario-based and time-based (temporal) decomposition ideas to efficiently solve the two-stage problem. Numerical experiments suggest that our approach significantly outperforms the CPLEX solver in both solution quality and computational time (see Section 6.2).

Based on a data set from Citi Bike in NYC, we investigate the roles of rider crowdsourcing and the 3PL in micromobility vehicle relocation for different parameter settings and demand patterns. We first solve the integrated vehicle allocation and relocation problem to obtain an initial vehicle allocation. Based on this initial allocation, we then solve the second-stage relocation problem to determine out-of-sample relocation decisions. We uncover the following insights that may shed light on the successful operations of shared micromobility systems.

(i) If the vehicle allocation budget is tight, the number of relocated vehicles in the second stage increases as the initial vehicle allocation increases. In contrast, if the vehicle allocation budget is sufficiently large, the number of relocated vehicles decreases as the initial vehicle allocation becomes larger. This is because if there are sufficiently many vehicles available initially, the need to relocate them to match supply with demand subsequently becomes smaller (see Section 6.3).

(ii) If both relocation methods are used, the number of vehicles relocated by the 3PL dominates that by rider crowdsourcing because of the 3PL's lower relocation cost per vehicle. The 3PL satisfies mostly mass relocation needs, while rider crowdsourcing mainly serves sporadic relocation needs. We find that investing in rider crowdsourcing can increase the profit, reduce the demand loss, and improve the vehicle utilization rate of the system with an existing 3PL relocation method. Therefore, in high-traffic areas, combining the two relocation methods leads to a higher profit.

(iii) As the rider crowdsourcing budget increases, the vehicle relocation by rider crowdsourcing increases, while the vehicle relocation by the 3PL first increases and then decreases. This suggests that rider crowdsourcing complements (benefits) the 3PL when the crowdsourcing budget is low, but substitutes the 3PL when the budget is high. Overall, introducing rider crowdsourcing to the micromobility system with 3PL relocation may significantly increase the system's profit without affecting the existing commitment with the 3PL (see Section 6.5).

(iv) Through studying the temporal and spatial relocation features, we find that the 3PL often conducts relocation around demand peak hours, while rider crowdsourcing is conducted throughout the day. To increase the supply in high-demand pick-up regions, the 3PL may relocate vehicles in batches from faraway, low-demand regions, while rider crowdsourcing tends to relocate a few vehicles each time from neighboring regions (see Section 6.6).

(v) The temporal demand pattern also has significant impact on vehicle allocation and relocation. The operator should allocate more vehicles initially under a bimodal customer arrival pattern. Besides, rider crowdsourcing relocates more vehicles under a unimodal customer arrival pattern than a bimodal pattern, whereas the reverse holds for the 3PL (see Section 6.7).

Our model and solution approach are sufficiently general for any typical shared micromobility systems. As the data from Citi Bike (2021) is quite representative, we believe the above insights shall hold for other micromobility systems with similar operational features as well. Our framework can also be potentially applied to car sharing systems where a single operator owns vehicles. Note that the 3PL uses only one truck with capacity $\bar{q}$ to perform its service in our paper. This assumption is supported by our numerical experiments. For all the periods and scenarios, the percentage of cases in which the load of the 3PL hits its capacity $\bar{q}$ is 0.84% on a weekday and 0.18% on a weekend. This is equivalent to $240 \times 0.84\% \approx 2$ periods on a weekday and $240 \times 0.18\% \approx 0.4$ periods on a weekend, which are negligible. The system operator can handle these cases separately when extra capacity is needed (e.g., by recruiting crowdsourced riders). In case one truck is not enough for the 3PL to cover the demand of the entire service area, we can split the service area into smaller parts and each part is served by one truck. We acknowledge this limitation of our model. We also simplify the cost structure of the shared micromobility system because of limited information available. The results may be improved with more information provided. In addition, besides the two temporal demand patterns studied, we can explore other demand distributions and customer behaviors to enrich the insights. We leave these for future research.

We believe that crowdsourcing can help address many other OM problems (Allon and Babich 2020) besides sustainable transportation. For example, independent car drivers (Qi et al. 2018, Fatehi and Wagner 2022) or cyclists (Kafle et al. 2017, Tu et al. 2019) are crowdsourced to perform last-mile delivery, labor is crowdsourced to contribute innovative ideas via completing a set of tasks on a virtual platform (Bayus 2013, Karger et al. 2014, Hu and Wang 2021), and volunteers are crowdsourced by non-profit organizations to complete tasks such as food recovery (Manshadi and Rodilitz 2022). More studies can be conducted to show the power of crowdsourcing.

Acknowledgments

The authors thank the Department Editor, Associate Editor, and two anonymous referees very much for their valuable comments and suggestions and thank Ding Zou for his early help. This work was supported by the Research Grants Council of Hong Kong [Grants 15501319 and 15505318]. Z. Jin was supported by the Hong Kong PhD Fellowship Scheme. Y. F. Lim was supported by the Lee Kong Chian School of Business, Singapore Management University [Maritime and Port Authority Research Fellowship] and the Key Program of the National Natural Science Foundation of China [Grant 71931009].

References

Allon G, Babich V (2020) Crowdsourcing and crowdfunding in the manufacturing and services sectors. *Manufacturing & Service Operations Management* 22(1):102–112.

Averbakh I, Berman O (1996) Locating flow-capturing units on a network with multi-counting and diminishing returns to scale. *European Journal of Operational Research* 91(3):495–506.

Bayus BL (2013) Crowdsourcing new product ideas over time: An analysis of the dell ideastorm community. *Management Science* 59(1):226–244.

Benjaafar S, Jiang D, Li X, Li X (2022) Dynamic inventory repositioning in on-demand rental networks. *Management Science* 68(11):7861–7878.

Bird (2020) *Vehicle Payouts*. Accessed: January 27, 2022.

Birge JR, Louveaux F (2011) *Introduction to Stochastic Programming* (Springer).

Boyacı B, Zografos KG, Geroliminis N (2015) An optimization framework for the development of efficient one-way car-sharing systems. *European Journal of Operational Research* 240(3):718–733.

Cannan E (1892) The origin of the law of diminishing returns, 1813-15. *The Economic Journal* 2(5):53–69.

Chang J, Yu M, Shen S, Xu M (2017) Location design and relocation of a mixed car-sharing fleet with a CO2 emission constraint. *Service Science* 9(3):205–218.

Cheng Y, Wang J, Wang Y (2021) A user-based bike rebalancing strategy for free-floating bike sharing systems: A bidding model. *Transportation Research Part E: Logistics and Transportation Review* 154:102438.

Chung YH, Färe R, Grosskopf S (1997) Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management* 51(3):229–240.

Citi Bike (2021) `https://citibikenyc.com/homepage`, accessed: January 27, 2022.

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

Fatehi S, Wagner MR (2022) Crowdsourcing last-mile deliveries. *Manuf. Serv. Oper. Manag.* 24(2):791–809.

Feng G, Kong G, Wang Z (2021) We are on the way: Analysis of on-demand ride-hailing systems. *Manufacturing & Service Operations Management* 23(5):1237–1256.

Freund D, Norouzi-Fard A, Paul A, Wang C, Henderson SG, Shmoys DB (2020) Data-driven rebalancing methods for bike-share systems. *Analytics for the Sharing Economy: Mathematics, Engineering and Business Perspectives*, 255–278 (Springer).

Fricker C, Gast N (2016) Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *EURO Journal on Transportation and Logistics* 5(3):261–291.

Fu C, Zhu N, Ma S, Liu R (2022) A two-stage robust approach to integrated station location and rebalancing vehicle service design in bike-sharing systems. *European Journal of Operational Research* 298(3):915–938.

Grahovac J, Chakravarty A (2001) Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. *Management Science* 47(4):579–594.

Hamilton JD (2020) *Time Series Analysis* (Princeton University Press).

Hasija S, Shen ZJM, Teo CP (2020) Smart city operations: Modeling challenges and opportunities. *Manufacturing & Service Operations Management* 22(1):203–213.

He L, Hu Z, Zhang M (2020) Robust repositioning for vehicle sharing. *Manufacturing & Service Operations Management* 22(2):241–256.

He L, Mak HY, Rong Y, Shen ZJM (2017) Service region design for urban electric vehicle sharing systems. *Manufacturing & Service Operations Management* 19(2):309–327.

He QC, Nie T, Yang Y, Shen ZJ (2021) Beyond repositioning: Crowd-sourcing and geo-fencing for shared-mobility systems. *Production and Operations Management* 30(10):3448–3466.

Homburg C, Ehm L, Artz M (2015) Measuring and managing consumer sentiment in an online community environment. *Journal of Marketing Research* 52(5):629–641.

Horwitz J (2017) *One of China's Top Bike-sharing Startups is Now Paying Users to Ride Its Bikes*. Accessed: January 27, 2022.

Hu M, Wang L (2021) Joint vs. separate crowdsourcing contests. *Management Science* 67(5):2711–2728.

Huang J, Chou MC, Teo CP (2021) Bike-repositioning using volunteers: Crowd sourcing with choice restriction. *Proc. of the AAAI Conference on Artificial Intelligence*, 11844–11852.

Kabra A, Belavina E, Girotra K (2020) Bike-share systems: Accessibility and availability. *Management Science* 66(9):3803–3824.

Kafle N, Zou B, Lin J (2017) Design and modeling of a crowdsource-enabled system for urban parcel relay and delivery. *Transportation Research Part B: Methodological* 99:62–82.

Karger DR, Oh S, Shah D (2014) Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research* 62(1):1–24.

Li Y, Liu Y (2021) The static bike rebalancing problem with optimal user incentives. *Transportation Research Part E: Logistics and Transportation Review* 146:102216.

Lu M, Chen Z, Shen S (2018) Optimizing the profitability and quality of service in carshare systems under demand uncertainty. *Manufacturing & Service Operations Management* 20(2):162–180.

Mak HY (2022) Enabling smarter cities with operations management. *Manufacturing & Service Operations Management* 24(1):24–39.

Manshadi V, Rodilitz S (2022) Online policies for efficient volunteer crowdsourcing. *Management Science* .

McKinsey (2019) *Micromobility's 15,000-mile Checkup*. Accessed: January 27, 2022.

McKinsey (2021) *Why micromobility is here to stay*. Accessed: January 27, 2022.

MTA (2020) *Subway and bus ridership for 2019*. Accessed: January 27, 2022.

Murray D, Glidewell S (2019) *An Analysis of the Operational Costs of Trucking: 2019 Update*. Accessed: January 27, 2022.

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

Nair R, Miller-Hooks E (2014) Equilibrium network design of shared-vehicle systems. *European Journal of Operational Research* 235(1):47–61.

NYC DOT (2019) Mobility Report. NYC Department of Transportation. Accessed: January 27, 2022.

Qi W, Li L, Liu S, Shen ZJM (2018) Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact. *Manuf. & Service Operations Management* 20(4):737–751.

Qi W, Shen ZJM (2019) A smart-city scope of operations management. *Prod. Oper. Manag.* 28(2):393–406.

Ravichandran T, Han S, Mithas S (2017) Mitigating diminishing returns to r&d: The role of information technology in innovation. *Information Systems Research* 28(4):812–827.

Rosenberg A (1992) *Economics–Mathematical Politics or Science of Diminishing Returns?* (University of Chicago Press).

Shephard RW, Färe R (1974) The law of diminishing returns. *Zeitschrift für Nationalökonomie* 34(1-2):69–90.

Shu J, Chou MC, Liu Q, Teo CP, Wang IL (2013) Models for effective deployment and redistribution of bicycles within public bicycle-sharing systems. *Operations Research* 61(6):1346–1359.

Simlett J, Møller TH (2020) *Micromobility: Moving Cities into a Sustainable Future*. Accessed: January 27, 2022.

Singla A, Santoni M, Bartók G, Mukerji P, Meenen M, Krause A (2015) Incentivizing users for balancing bike sharing systems. *Proc. of the AAAI Conference on Artificial Intelligence*, 723–729.

Srinivasan R, Moorman C (2005) Strategic firm commitments and rewards for customer relationship management in online retailing. *Journal of Marketing* 69(4):193–200.

Stokkink P, Geroliminis N (2021) Predictive user-based relocation through incentives in one-way carsharing systems. *Transportation Research Part B: Methodological* 149:230–249.

Tagaras G, Cohen MA (1992) Pooling in two-location inventory systems with non-negligible replenishment lead times. *Management Science* 38(8):1067–1083.

Tu W, Zhao T, Zhou B, Jiang J, Xia J, Li Q (2019) OCD: Online crowdsourced delivery for on-demand food. *IEEE Internet of Things Journal* 7(8):6842–6854.

Online Supplement for "Z. Jin, Y. Wang, Y. F. Lim, K. Pan*, and Z.-J. M. Shen. Vehicle Rebalancing in A Shared Micromobility System with Rider Crowdsourcing. ***Manufacturing and Service Operations Management*****."**

Appendix A: Table of Notation

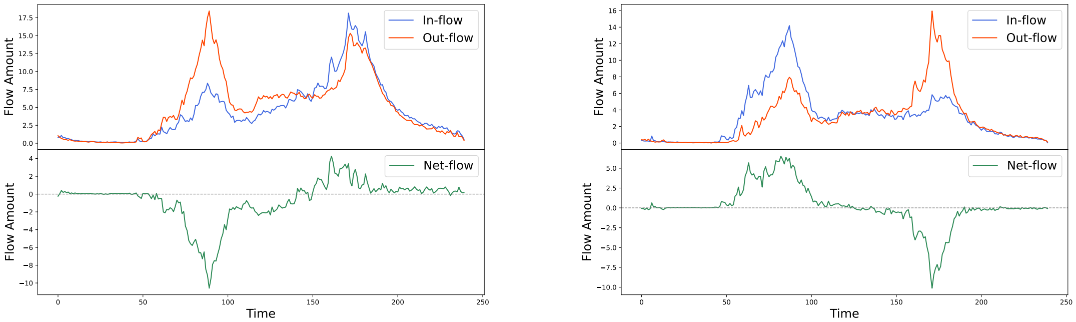
Table A1 Summary of Notation

Notation	Description
Sets:	
$\mathcal{V}$	set of service regions $\mathcal{V} = \{1, 2, \ldots, V\}$
$\mathcal{T}$	operational horizon $\mathcal{T} = \{0, 1, \ldots, T-1\}$
$\mathcal{T}(l,s,e)$	$\{s, s+1, \ldots, e-l-1\}$
$\mathcal{G}$	time-space network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{A}$ is the set of directed arcs
$\mathcal{A}^{\rm t}, \mathcal{A}^{\rm i}, \mathcal{A}^{\rm r}$	trip arcs, idle arcs, 3PL relocation arcs
$\mathcal{K}$	set of scenarios of uncertain demands in all the service regions across all the periods.
$\mathcal{H}$	set of segments on the interval $[0, \bar{\phi}]$ used to approximate the incentive function $g(\cdot)$, $\mathcal{H} = \{1, 2, \ldots, H\}$
Parameters:	
l_{ij}	duration for a rider trip from region $i \in \mathcal{V}$ to region $j \in \mathcal{V}$
$l_{\rm r}$	duration for the 3PL to relocate vehicles from region $i \in \mathcal{V}$ to region $j \in \mathcal{V}$
N	upper bound for the total number of allocated vehicles
B_j	upper bound for the number of vehicles allocated to region $j \in \mathcal{V}$
c_j	cost incurred for allocating a micromobility vehicle to region $j \in \mathcal{V}$
λ_a	customer demand (the number of customers) from region i in period t to region j with travel time l_{ij} for each arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\rm t}$
R, $C_{\rm p}$	revenue for serving a customer per period and penalty cost per customer lost, respectively
$\bar{\Lambda}_{ij}$	maximum number of riders that can be crowdsourced from region i to conduct trips to region j
β_{ij}	the rate of diminishing return on rewards in the incentive function $g(\cdot)$
$B_{\rm c}$	upper bound for the total incentive used by the operator to crowdsource riders
$C_{\rm r}$	fixed fee paid to the 3PL for relocating vehicles per request
$\bar{z}$	maximum number of times that the 3PL can relocate vehicles per day
$\bar{\bar{z}}$	maximum number of periods that the 3PL operates during any time interval of $l_{\rm f}$ periods
$\underline{q}$, $\bar{q}$	lower and upper bounds for the total number of relocated vehicles by the 3PL in each period
p^k	probability of each scenario $k \in \mathcal{K}$
$\bar{\phi}$	Given an arbitrarily small number $\epsilon > 0$, $\bar{\phi} = g^{-1}(\bar{\Lambda} - \epsilon)$
$\bar{k}_h$	slope of a linear function passing through points $((h-1)\delta, g((h-1)\delta))$ and $(h\delta, g(h\delta))$, $\bar{k}_h = (g(h\delta) - g((h-1)\delta))/\delta$ with $\delta = \bar{\phi}/H$
M	number of sub-networks for temporal decomposition
$T_{\rm sub}$	number of periods in each sub-network $m \in \{1, 2, \ldots, M\}$
$\mathcal{D}_m$	dictionary to record the pair of initial vehicle allocation and the corresponding objective value of the two-stage problem over sub-network m
$\bar{N}_1, \bar{N}_2$	step sizes for heuristic search
Variables:	
x_i	number of vehicle allocated to region $i \in \mathcal{V}$
$\tilde{x}_i$	number of vehicles in the ending period $T-1$ in region $i \in \mathcal{V}$
η_a	unsatisfied demand on arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\rm t}$
Λ_a	number of crowdsourced riders from region i in period t to region j for a given arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\rm t}$
ϕ_a	incentive used to motivate the riders to relocate the vehicles along arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\rm t}$
y_a	realized flow on arc $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\rm i} \cup \mathcal{A}^{\rm t}$
γ_a	number of vehicles relocated by the 3PL from region i in period t to region j for each arc $a = (n_{it}, n_{j,t+l_{\rm r}}) \in \mathcal{A}^{\rm r}$
z_t	$z_t = 1$ if the 3PL is requested by the operator in period t to relocate vehicles and $z_t = 0$ otherwise
$\widetilde{z}_t$	the status of the 3PL such that $\widetilde{z}_t = 1$ if the 3PL provides service in period t and $\widetilde{z}_t = 0$ otherwise
u_h	u_h = 1 if the optimal reward ϕ falls in the hth segment for each $h \in \mathcal{H}$, and $u_h = 0$ otherwise
$\widetilde{\phi}_h$	$\widetilde{\phi}_h \in [0, \delta]$ such that the optimal reward $\phi = (h-1)\delta + \widetilde{\phi}_h$ for each $h \in \mathcal{H}$
ω_h	variable to approximate $u_h \widetilde{\phi}_h$ for each segment $h \in \mathcal{H}$

Appendix B: Data Analyses Showing Vehicle Supply-Demand Imbalance

Since (i) the customers use the micromobility vehicles for mostly one-way trips, (ii) many customers have similar travel patterns, and (iii) the demands may be uncertain with respect to time and space, the vehicles are depleted in some regions but are cluttered in the others. This leads to an imbalance of vehicle supply and demand. We use the data from Citi Bike (2021) in New York City (NYC) to illustrate this imbalance.

We track the numbers of bikes arriving at and leaving from each region, corresponding to its *in-flow* and *out-flow* respectively, in each period. For each period, define *net-flow* as the in-flow minus the out-flow. Intuitively, a zero net-flow in a region reflects a balance of vehicle supply and demand. Otherwise, a positive net-flow indicates an oversupply of vehicles and a negative net-flow may cause a shortage. Figure 16(a) shows a negative net-flow for an extended duration in the morning and a positive net-flow for an extended duration in the afternoon in region 5. In contrast, Figure 16(b) shows an opposite phenomenon in region 9: a consistent positive net-flow in the morning and a consistent negative net-flow in the afternoon. By comparing the net-flow of these two regions, we observe: (i) In the morning, the net-flow of bikes leaving from region 5 is larger than the net-flow of bikes going to region 9. (ii) In the evening, the net-flow going to region 5 is smaller than the net-flow leaving from region 9. This example demonstrates that an imbalance of supply and demand can occur in opposite manners simultaneously in different regions of a service area. This motivates the operator to rebalance the vehicles across different regions over the operational horizon.

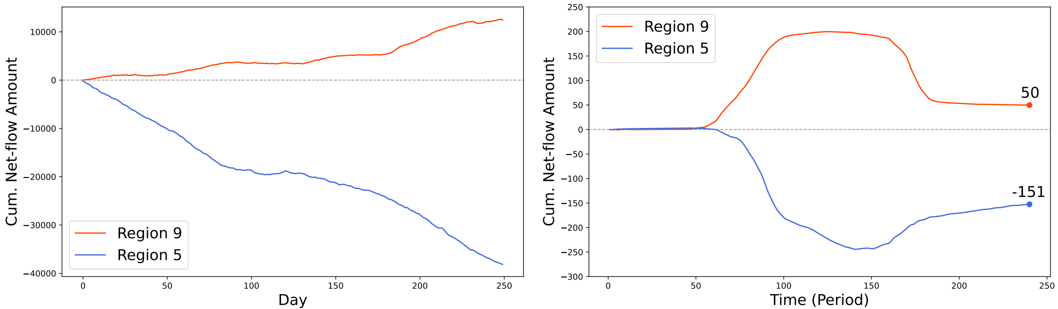


(a) Region 5

(b) Region 9

Figure 16 Supply-Demand Imbalance at the Region Level

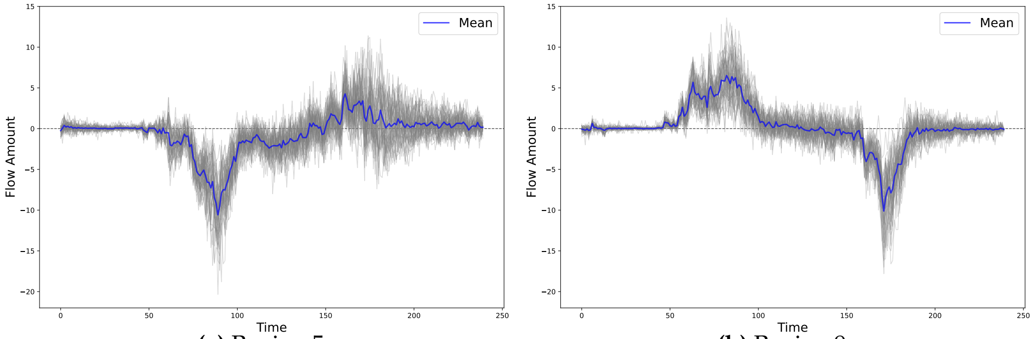
Figure 17(a) shows the cumulative net-flow of each region over a year. The figure shows that the cumulative net-flow of region 5 is always negative, and it has an overall decreasing trend. This suggests that region 5 is loosing bikes over time and may eventually face shortages. In contrast, the cumulative net-flow of region 9 is always positive, and it has an overall increasing trend. This suggests that region 9 is gaining bikes over time and may have an oversupply of bikes. Figure 17(b) shows the cumulative net-flow of each region within a day. Both Figures 17(a) and (b) show that the loss of bikes in region 5 is greater than the growth of bikes in region 9, which implies that there are bikes traveling out from these two regions and not coming back. This demonstrates that these two regions cannot naturally balance themselves.



(a) Over a Year

(b) Within a Day

Figure 17 Cumulative Net-flow



(a) Region 5

(b) Region 9

Figure 18 Net-flow over a Day

Therefore, this suffices to show that if the demand for bikes in each region is recurrent and consistent every day, then region 5 will run out of bikes and incur lost sales at some point in time without vehicle relocation. Figure 18 shows the net-flow of each region over a day. The gray curves correspond to 50 sample paths from the data of Citi Bike (2021). The blue curve represents the mean net-flow of all the sample paths in the data set. Both the gray and the blue curves suggest that the demand pattern of each region over the periods is quite consistent. Thus, if we do not relocate bikes from other regions to region 5, we can foresee that region 5 will run out of bikes and incur lost sales.

Finally, our numerical results in Section 6.4 also demonstrate the existence of vehicle supply-demand imbalance. Table 2 shows that compared to a system without any relocation operations, *introducing relocation operations increases the total profit, reduces demand loss, and raises vehicle utilization significantly*. Furthermore, Section 6.6 suggests that the optimal relocation by the 3PL and crowdsourcing is mostly performed during demand-peak hours, in which the supply-demand imbalance is most severe. This clearly indicates a need to address the vehicle supply-demand imbalance, and there is a significant value to solve the problem that occurs in the real world.

Appendix C: Supplement to Section 3

C.1. Undefined Arcs

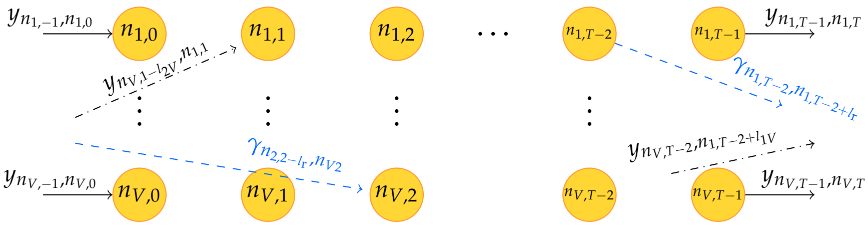


Figure 19 The Example of Undefined Arcs

C.2. Vector Definition

For each $k \in \mathcal{K}$, we let $\boldsymbol{\lambda}^k$ denote the demand vector and define $\widetilde{\mathbf{x}}^k = (\widetilde{x}_j^k,\ j \in \mathcal{V})^\top$ as the idle vehicle flows after period $T-1$. Similarly, for each $k \in \mathcal{K}$, we let $\boldsymbol{\eta}^k = (\eta_a^k,\ a \in \mathcal{A}^{\mathrm{t}})^\top$ represent the vector of unsatisfied demands, $\mathbf{y}^k = (y_a^k,\ a \in \mathcal{A}^{\mathrm{t}})^\top$ the vector of realized flows, $\boldsymbol{\gamma}^k = (\gamma_a^k,\ a \in \mathcal{A}^{\mathrm{r}})^\top$ the vector of 3PL relocation flows, $\boldsymbol{\Lambda}^k = (\Lambda_a^k,\ a \in \mathcal{A}^{\mathrm{t}})^\top$ the vector of rider crowdsourcing relocation flows, $\boldsymbol{\phi}^k = (\phi_a^k,\ a \in \mathcal{A}^{\mathrm{t}})^\top$ the vector of rewards to crowdsource riders, $\mathbf{z}^k = (z_t^k,\ t \in \mathcal{T}(l_{\mathrm{r}}))^\top$ the vector of 3PL relocation requests, and $\widetilde{\mathbf{z}}^k = (\widetilde{z}_t^k,\ t \in \mathcal{T}(l_{\mathrm{r}}))^\top$ the vector of the 3PL relocation operation status (i.e., whether or not it is active). Moreover, for each $k \in \mathcal{K}$, let $\mathbf{v}^k = (\widetilde{\mathbf{x}}^k, \boldsymbol{\eta}^k, \mathbf{y}^k, \boldsymbol{\gamma}^k)$. We let $\widetilde{\mathbf{x}} = (\widetilde{\mathbf{x}}^k,\ t \in \mathcal{K})^\top$, $\boldsymbol{\eta} = (\boldsymbol{\eta}^k,\ t \in \mathcal{K})^\top$, $\mathbf{y} = (\mathbf{y}^k,\ t \in \mathcal{K})^\top$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}^k,\ t \in \mathcal{K})^\top$, $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}^k,\ t \in \mathcal{K})^\top$, $\boldsymbol{\phi} = (\boldsymbol{\phi}^k,\ t \in \mathcal{K})^\top$, $\mathbf{z} = (\mathbf{z}^k,\ t \in \mathcal{K})^\top$, and $\widetilde{\mathbf{z}} = (\widetilde{\mathbf{z}}^k,\ t \in \mathcal{K})^\top$. In addition, we let $\mathbf{c} = (c_i,\ i \in \mathcal{V})^\top$ and $\boldsymbol{\rho} = (\mathbf{x}, \tilde{\mathbf{z}}, \mathbf{z}, \mathbf{v}, \boldsymbol{\Lambda})$.

Appendix D: Supplement to Section 4

D.1. Constraints

The constraint set $\mathcal{Y}(x_1, x_2, \lambda)$ is defined by the following constraints:

$$y_{n_{it},n_{i,t+1}} + y_{n_{it},n_{3-i,t+L}} + \gamma_{n_{it},n_{3-i,t+L}} = \begin{cases} x_i,\ t=0, \\ y_{n_{i,t-1},n_{it}},\ t=1,2,\ldots,L-1, \\ y_{n_{i,t-1},n_{it}} + y_{n_{3-i,t-L},n_{it}} + \gamma_{n_{3-i,t-L},n_{it}},\ t=L,\ldots,T-1, \end{cases} \quad i \in \{1,2\}, \tag{8a}$$

$$y_{n_{it},n_{3-i,t+L}} = \begin{cases} \lambda_{n_{it},n_{3-i,t+L}} + \Lambda_{n_{it},n_{3-i,t+L}} - \eta_{n_{it},n_{3-i,t+L}},\ t=0,1,\ldots,T-L-1, \\ 0,\ t=T-L,\ldots,T-1, \end{cases} \quad i \in \{1,2\}, \tag{8b}$$

$$\gamma_{n_{it},n_{3-i,t+L}} \le \bar{q} z_t,\ i \in \{1,2\},\ t \in \{0,1,\ldots,T-L-1\}, \tag{8c}$$

$$\begin{aligned} &\Lambda_{n_{it},n_{3-i,t+L}},\ y_{n_{it},n_{3-i,t+L}},\ \eta_{n_{it},n_{3-i,t+L}}, \gamma_{n_{it},n_{3-i,t+L}} \ge 0,\ i \in \{1,2\},\ t \in \{0,1,\ldots,T-L-1\}, \\ &z_t \in \{0,1\},\ t \in \{0,1,\ldots,T-L-1\};\ y_{n_{it},n_{i,t+1}} \ge 0,\ i \in \{1,2\},\ t \in \mathcal{T} \setminus \{T-1\}. \end{aligned} \tag{8d}$$

D.2. Proof of Proposition 1

Proof. By substituting constraints (8b) to the objective for all $t \in \{0,1,\ldots,T-L-1\}$ and removing the constant term, we can equivalently replace the objective in (4) by the following one:

$$\sum_{t=0}^{T-L-1}\left(\sum_{i=1}^{2}\Big((C_\mathrm{p}+RL)\,\eta_{n_{it},n_{3-i,t+L}}+\alpha\Lambda_{n_{it},n_{3-i,t+L}}\Big)+C_\mathrm{r}z_t\right). \tag{9}$$

We denote the optimal solution of problem (4) by $\mathbf{Y}^*=(\Lambda^*_{n_{it},n_{3-i,t+L}},\ y^*_{n_{it},n_{3-i,t+L}},\ \eta^*_{n_{it},n_{3-i,t+L}},\ \gamma^*_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-L-1\},\ y^*_{n_{it},n_{i,t+1}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-1\},\ z^*_t,\ t\in\{0,1,\ldots,T-L-1\})$. We then prove the proposition by contradiction and we first focus on Part (b) and Part (c) and then Part (a).

For Part (b), suppose, on the contrary, that there exists some $\tilde{i}\in\{1,2\}$ or $\tilde{t}\in\{0,1,\ldots,T-L-1\}$ such that one of the following three cases happen: (i) $\alpha\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}>C_\mathrm{r}$, (ii) $\alpha\Lambda^*_{n_{\tilde{i}t},n_{3-\tilde{i},t+L}}>C_\mathrm{r}$ for any $t\in\{0,1,\ldots,T-L-1\}$, and (iii) $\alpha\Lambda^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}>C_\mathrm{r}$ for any $i\in\{1,2\}$. For case (i), we construct a new solution $\mathbf{Y}^\dagger=(\Lambda^\dagger_{n_{it},n_{3-i,t+L}},\ y^\dagger_{n_{it},n_{3-i,t+L}},\ \eta^\dagger_{n_{it},n_{3-i,t+L}},\ \gamma^\dagger_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-L-1\},\ y^\dagger_{n_{it},n_{i,t+1}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-1\},\ z^\dagger_t,\ t\in\{0,1,\ldots,T-L-1\})$ such that

$$\Lambda^\dagger_{n_{\tilde{i},\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=0,\ \gamma^\dagger_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}+\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}},\ z^\dagger_{\tilde{t}}=1,\ y^\dagger_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=y^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}-\Lambda^*_{n_{\tilde{i},\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}, \tag{10a}$$

$$\Lambda^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=\Lambda^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ y^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=y^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ \gamma^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=\gamma^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ i\in\{1,2\}\backslash\{\tilde{i}\}, \tag{10b}$$

$$\Lambda^\dagger_{n_{it},n_{3-i,t+L}}=\Lambda^*_{n_{it},n_{3-i,t+L}},\ y^\dagger_{n_{it},n_{3-i,t+L}}=y^*_{n_{it},n_{3-i,t+L}},\ \gamma^\dagger_{n_{it},n_{3-i,t+L}}=\gamma^*_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},t\in\{0,1,\ldots,T-L-1\}\backslash\{\tilde{t}\}, \tag{10c}$$

$$z^\dagger_t=z^*_t,\ t\in\{0,1,\ldots,T-L-1\}\backslash\{\tilde{t}\};\ \eta^\dagger_{n_{it},n_{3-i,t+L}}=\eta^*_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},t\in\{0,1,\ldots,T-L-1\}, \tag{10d}$$

$$y^\dagger_{n_{it},n_{i,t+1}}=y^*_{n_{it},n_{i,t+1}},\ i\in\{1,2\},t\in\mathcal{T}. \tag{10e}$$

It is easy to verify that $\mathbf{Y}^\dagger$ is feasible for problem (4). It follows that the objective (9) with respect to $\mathbf{Y}^\dagger$ minus that with respect to $\mathbf{Y}^*$ equals $-\alpha\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}+C_\mathrm{r}(1-z^*_{\tilde{t}})\le-\alpha\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}+C_\mathrm{r}<0$, where the first inequality holds because $C_\mathrm{r}\ge 0$ and $z^*_{\tilde{t}}\ge 0$ and the second inequality holds because of the case (i) assumption. Thus, a lower objective is achieved with respect to the solution $\mathbf{Y}^\dagger$, contradicting that $\mathbf{Y}^*$ is the optimal solution. For both cases (ii) and (iii), we can similarly construct such a new solution $\mathbf{Y}^\dagger$ and lead to the contradiction. We omit the details for brevity. Therefore, $\alpha\Lambda^*_{n_{it},n_{3-i,t+L}}\le C_\mathrm{r}$ holds for any $i\in\{1,2\}$ and $t\in\{0,1,\ldots,T-L-1\}$.

For Part (c), suppose, on the contrary, that there exists some $\tilde{i}\in\{1,2\}$ or $\tilde{t}\in\{0,1,\ldots,T-L-1\}$ such that one of the following cases happen: (i) $C_\mathrm{r}>\alpha\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}$ if $z^*_{\tilde{t}}=1$ and $\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}>0$, (ii) $C_\mathrm{r}>\alpha\gamma^*_{n_{\tilde{i}t},n_{3-\tilde{i},t+L}}$ if $z^*_t=1$ and $\gamma^*_{n_{\tilde{i}t},n_{3-\tilde{i},t+L}}>0$ for any $t\in\{0,1,\ldots,T-L-1\}$, and (iii) $C_\mathrm{r}>\alpha\gamma^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}$ if $z^*_{\tilde{t}}=1$ and $\gamma^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}>0$ for any $i\in\{1,2\}$. For case (i), we construct a new solution $\mathbf{Y}^\dagger=(\Lambda^\dagger_{n_{it},n_{3-i,t+L}},\ y^\dagger_{n_{it},n_{3-i,t+L}},\ \eta^\dagger_{n_{it},n_{3-i,t+L}},\ \gamma^\dagger_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-L-1\},\ y^\dagger_{n_{it},n_{i,t+1}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-1\},\ z^\dagger_t,\ t\in\{0,1,\ldots,T-L-1\})$ such that

$$\Lambda^\dagger_{n_{\tilde{i},\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}, \tag{11a}$$

$$\gamma^\dagger_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=0,\ z^\dagger_{\tilde{t}}=0,\ y^\dagger_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}=y^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}+\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}, \tag{11b}$$

$$\Lambda^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=\Lambda^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ y^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=y^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ \gamma^\dagger_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}=\gamma^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}},\ i\in\{1,2\}\backslash\{\tilde{i}\}, \tag{11c}$$

$$\Lambda^\dagger_{n_{it},n_{3-i,t+L}}=\Lambda^*_{n_{it},n_{3-i,t+L}},\ y^\dagger_{n_{it},n_{3-i,t+L}}=y^*_{n_{it},n_{3-i,t+L}},\ \gamma^\dagger_{n_{it},n_{3-i,t+L}}=\gamma^*_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},t\in\{0,1,\ldots,T-L-1\}\backslash\{\tilde{t}\}, \tag{11d}$$

$$z^\dagger_t=z^*_t,\ t\in\{0,1,\ldots,T-L-1\}\backslash\{\tilde{t}\};\ \eta^\dagger_{n_{it},n_{3-i,t+L}}=\eta^*_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},t\in\{0,1,\ldots,T-L-1\}, \tag{11e}$$

$$y^\dagger_{n_{it},n_{i,t+1}}=y^*_{n_{it},n_{i,t+1}},\ i\in\{1,2\},t\in\mathcal{T}. \tag{11f}$$

It is easy to verify that $\mathbf{Y}^\dagger$ is feasible for problem (4). Then, the objective (9) with respect to $\mathbf{Y}^\dagger$ minus that with respect to $\mathbf{Y}^*$ equals $\alpha\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}-C_\mathrm{r}<0$ due to the case (i) assumption. Thus, a lower objective is achieved with respect to $\mathbf{Y}^\dagger$, contradicting that $\mathbf{Y}^*$ is the optimal solution. For both cases (ii) and (iii), we can similarly construct such a new solution $\mathbf{Y}^\dagger$ and lead to the contradiction. We omit the details for brevity. Thus, $C_\mathrm{r}\le\alpha\gamma^*_{n_{it},n_{3-i,t+L}}$ holds for any $i\in\{1,2\}$ and $t\in\{0,1,\ldots,T-L-1\}$ if $z^*_t=1$ and $\gamma^*_{n_{it},n_{3-i,t+L}}>0$.

For Part (a), suppose, on the contrary, that there exists some $\tilde{i}\in\{1,2\}$ or $\tilde{t}\in\{0,1,\ldots,T-L-1\}$ such that one of the following cases happen: (i) $\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}>0$ and $\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}>0$, (ii) $\Lambda^*_{n_{\tilde{i}t},n_{3-\tilde{i},t+L}}>0$ and $\gamma^*_{n_{\tilde{i}t},n_{3-\tilde{i},t+L}}>0$ for any $t\in\{0,1,\ldots,T-L-1\}$, and (iii) $\Lambda^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}>0$ and $\gamma^*_{n_{i\tilde{t}},n_{3-i,\tilde{t}+L}}>0$ for any $i\in\{1,2\}$. For case (i), we further consider two subcases. Subcase 1): $\Lambda^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}+\gamma^*_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}\le C_\mathrm{r}/\alpha$. We construct a new solution $\mathbf{Y}^\dagger=(\Lambda^\dagger_{n_{it},n_{3-i,t+L}},\ y^\dagger_{n_{it},n_{3-i,t+L}},\ \eta^\dagger_{n_{it},n_{3-i,t+L}},\ \gamma^\dagger_{n_{it},n_{3-i,t+L}},\ i\in\{1,2\},\ t\in\{0,1,\ldots,T-L-1\},\ y^\dagger_{n_{it},n_{i,t+1}},\ i\in$

$\{1,2\}$, $t \in \{0,1,\ldots,T-1\}$, $z_t^{\dagger}$, $t \in \{0,1,\ldots,T-L-1\}$) such that $\Lambda^{\dagger}_{n_{\tilde{i},\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}} = \Lambda^{*}_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}} + \gamma^{*}_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}}$ and (11b)–(11f) hold. Subcase 2): $\Lambda^{*}_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}} + \gamma^{*}_{n_{\tilde{i}\tilde{t}},n_{3-\tilde{i},\tilde{t}+L}} > C_{\mathrm{r}}/\alpha$. We construct a new solution $\mathbf{Y}^{\dagger}$ such that (10a)–(10e) hold. The objective (9) with respect to $\mathbf{Y}^{\dagger}$ in either subcase is strictly lower as shown before, which contradicts that $\mathbf{Y}^{*}$ is the optimal solution of problem (4). For both cases (ii) and (iii), we can similarly construct such a new solution $\mathbf{Y}^{\dagger}$ and lead to the contradiction. We omit the details for brevity. Therefore, $\Lambda^{*}_{n_{it},n_{3-i,t+L}}\gamma^{*}_{n_{it},n_{3-i,t+L}} = 0$ holds for any $i \in \{1,2\}$ and $t \in \{0,1,\ldots,T-L-1\}$. □

Appendix E: Supplement to Section 5

E.1. Proof of Proposition 2

Proof. Note that given the allocation solution $\mathbf{x}$, the solution to the second-stage problem ($\mathcal{Q}_0$) (or ($\mathcal{Q}$)) with respect to any scenario $k \in \mathcal{K}$ is independent from that with respect to any other scenario $k' \in \mathcal{K}$. Thus, it suffices to consider only one scenario in $\mathcal{K}$ for both problems ($\mathcal{Q}_0$) and ($\mathcal{Q}$). To that end, we omit the superscript k in both problems ($\mathcal{Q}_0$) and ($\mathcal{Q}$) and denote by $\widetilde{\mathbf{Y}} = (\widetilde{\mathbf{x}}, \boldsymbol{\eta}, \mathbf{y}, \gamma, \mathbf{z}, \widetilde{\mathbf{z}}, \boldsymbol{\Lambda}, \boldsymbol{\phi}, \mathbf{u}, \widetilde{\boldsymbol{\phi}}, \boldsymbol{\omega})$ the second-stage variables of problem ($\mathcal{Q}_0$) or ($\mathcal{Q}$) with respect to any scenario in $\mathcal{K}$. For ease of notation, we also split $\widetilde{\mathbf{Y}}$ into three parts: $\widetilde{\mathbf{Y}}_1 = (\widetilde{\mathbf{x}}, \boldsymbol{\eta}, \mathbf{y}, \gamma, \mathbf{z}, \widetilde{\mathbf{z}})$, $\widetilde{\mathbf{Y}}_2 = (\boldsymbol{\Lambda}, \boldsymbol{\phi})$ and $\widetilde{\mathbf{Y}}_3 = (\mathbf{u}, \widetilde{\boldsymbol{\phi}}, \boldsymbol{\omega})$. Moreover, to differentiate the notation in problems ($\mathcal{Q}_0$) and ($\mathcal{Q}$), we denote the variables of problem ($\mathcal{Q}_0$) by $\widetilde{\mathbf{Y}}^{\mathrm{b}}$ and the corresponding feasible region by $\mathcal{Y}^{\mathrm{b}} = \{\widetilde{\mathbf{Y}}^{\mathrm{b}} \mid (2a) - (2h), (1b), (1c), (3), ((5c) - (5e), (6),\ a \in \mathcal{A}^{\mathrm{t}})\}$. Similarly, we denote the variables of problem ($\mathcal{Q}$) by $\widetilde{\mathbf{Y}}^{\mathrm{r}}$ and the corresponding feasible region by $\mathcal{Y}^{\mathrm{r}} = \{\widetilde{\mathbf{Y}}^{\mathrm{r}} \mid (2a) - (2h), (1b), (1c), (3), ((5c) - (5d), (6), (u_{h,a} \in [0,1],\ h \in \mathcal{H}),\ a \in \mathcal{A}^{\mathrm{t}})\}$. The optimal solutions of both ($\mathcal{Q}_0$) and ($\mathcal{Q}$) are specified with superscript $\cdot^{*}$. The optimal objective values of both problems are denoted by $\Theta'(\mathbf{x}) = \Psi(\widetilde{\mathbf{Y}}^{\mathrm{b}*})$ and $\Theta''(\mathbf{x}) = \Psi(\widetilde{\mathbf{Y}}^{\mathrm{r}*})$. Because $\mathcal{Y}^{\mathrm{b}} \subseteq \mathcal{Y}^{\mathrm{r}}$ and both problems ($\mathcal{Q}_0$) and ($\mathcal{Q}$) are minimization problems sharing the same objective function (7), we have

$$\Theta'(\mathbf{x}) \geq \Theta''(\mathbf{x}). \tag{12}$$

In the following, we will show that $\Theta'(\mathbf{x}) \leq \Theta''(\mathbf{x})$. We claim that

$$\begin{aligned} &\left(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_2^{\mathrm{r}*}\right) \in \mathrm{Proj}_{\left(\widetilde{\mathbf{Y}}_1^{\mathrm{b}}, \widetilde{\mathbf{Y}}_2^{\mathrm{b}}\right)}\left(\mathcal{Y}^{\mathrm{b}}\right) \\ &= \left\{\left(\widetilde{\mathbf{Y}}_1^{\mathrm{b}}, \widetilde{\mathbf{Y}}_2^{\mathrm{b}}\right) \in \mathbb{R}_+^{(|\mathcal{V}|+4|\mathcal{A}^{\mathrm{t}}|+|\mathcal{A}^{\mathrm{r}}|+2|\mathcal{T}(l_{\mathrm{r}})|)} \;\middle|\; \exists \widetilde{\mathbf{Y}}_3^{\mathrm{b}} \in \mathbb{R}_+^{3H|\mathcal{A}^{\mathrm{t}}|} \text{ such that } \widetilde{\mathbf{Y}}^{\mathrm{b}} \in \mathcal{Y}^{\mathrm{b}}\right\}. \end{aligned} \tag{13}$$

Note that $(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_2^{\mathrm{r}*})$ satisfies constraints (2a)–(2h), (1b), (1c), and (3) because these constraints are shared by $\mathcal{Y}^{\mathrm{b}}$ and $\mathcal{Y}^{\mathrm{r}}$ and $\widetilde{\mathbf{Y}}_3$ is involved in none of these constraints. Thus, to show that our claim holds, it suffices to show that there exists $\widetilde{\mathbf{Y}}_3^{\mathrm{r}'} \in \mathbb{R}_+^{3H|\mathcal{A}^{\mathrm{t}}|}$ such that $(\widetilde{\mathbf{Y}}_2^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_3^{\mathrm{r}'})$ satisfies (5c)–(5e) and (6) for any $a \in \mathcal{A}^{\mathrm{t}}$, thereby leading to $(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_2^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_3^{\mathrm{r}'}) \in \mathcal{Y}^{\mathrm{b}}$. We will prove this by contradiction.

Note that for any given $a \in \mathcal{A}^{\mathrm{t}}$, if $(\widetilde{\mathbf{Y}}_2^{\mathrm{r}*}, \widetilde{\mathbf{Y}}_3^{\mathrm{r}'})$ satisfies (5c)–(5e) and (6), then $\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{r}*} = (\phi^{\mathrm{r}*}, \Lambda^{\mathrm{r}*})$ amounts to a point in the projection space $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}}}(\mathcal{Y}^{\mathrm{b}}) = \{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}} \in \mathbb{R}_+^2 \mid \exists(\widetilde{\mathbf{Y}}_1^{\mathrm{b}}, \widetilde{\mathbf{Y}}_3^{\mathrm{b}}) \text{ and } (\widetilde{\mathbf{Y}}_{2,e}^{\mathrm{b}}, e \in \mathcal{A}^{\mathrm{t}}, e \neq a) \text{ such that } \widetilde{\mathbf{Y}}^{\mathrm{b}} \in \mathcal{Y}^{\mathrm{b}}\}$. This projection space can be represented by a piece-wise linear curve with H segment and ϕ runs in $[0, \bar{\phi}]$, as illustrated in Figure 20, and such a point can be illustrated by point B in this figure. Now, for the contradiction, we suppose that there exists some $a' \in \mathcal{A}^{\mathrm{t}}$ such that $\widetilde{\mathbf{Y}}_{2,a'}^{\mathrm{r}*} = (\phi_{a'}^{\mathrm{r}*}, \Lambda_{a'}^{\mathrm{r}*})$ does not amount to a point in the projection space $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}}}(\mathcal{Y}^{\mathrm{b}})$. More specifically, by the feasibility of $\widetilde{\mathbf{Y}}_{2,a'}^{\mathrm{r}*}$ from ($\mathcal{Q}$), it will amount to a point in the projection space $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{r}}}(\mathcal{Y}^{\mathrm{r}}) = \{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{r}} \in \mathbb{R}_+^2 \mid \exists(\widetilde{\mathbf{Y}}_1^{\mathrm{r}}, \widetilde{\mathbf{Y}}_3^{\mathrm{r}}) \text{ and } (\widetilde{\mathbf{Y}}_{2,e}^{\mathrm{r}}, e \in \mathcal{A}^{\mathrm{t}}, e \neq a) \text{ such that } \widetilde{\mathbf{Y}}^{\mathrm{r}} \in \mathcal{Y}^{\mathrm{r}}\}$, which is the convex hull of $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}}}(\mathcal{Y}^{\mathrm{b}})$ and can be illustrated by the gray region Ω in Figure 20. Also, $\widetilde{\mathbf{Y}}_{2,a'}^{\mathrm{r}*}$ can be illustrated by point P.

Now, we construct a solution $(\widetilde{\mathbf{Y}}_1^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_2^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_3^{\mathrm{b}\prime})$. We let $\widetilde{\mathbf{Y}}_1^{\mathrm{b}\prime} = \widetilde{\mathbf{Y}}_1^{\mathrm{r}*}$. For $\widetilde{\mathbf{Y}}_2^{\mathrm{b}\prime} = (\phi^{\mathrm{b}\prime}, \Lambda^{\mathrm{b}\prime})$, we let $(\phi_e^{\mathrm{b}\prime}, \Lambda_e^{\mathrm{b}\prime}) = (\phi_e^{\mathrm{r}*}, \Lambda_e^{\mathrm{r}*})$ for any $e \in \mathcal{A}^{\mathrm{t}}$ and $e \neq a'$, $\Lambda_{a'}^{\mathrm{b}\prime} = \Lambda_{a'}^{r*}$, and $\phi_{a'}^{\mathrm{b}\prime} = (h'-1)\delta + (\Lambda^{\mathrm{r}*} - \sum_{m=1}^{h'-1} \bar{k}_m \delta)/\bar{k}_{h'}$, where $h' = \arg\min\{h \in \{1,2,\ldots,H\} | \sum_{m=1}^{h-1} \bar{k}_m \delta \leq \Lambda_{a'}^{\mathrm{r}*} < \sum_{m=1}^{h} \bar{k}_m \delta\}$. For $\widetilde{\mathbf{Y}}_3^{\mathrm{b}\prime} = (\mathbf{u}^{\mathrm{b}\prime}, \widetilde{\boldsymbol{\phi}}^{\mathrm{b}\prime}, \boldsymbol{\omega}^{\mathrm{b}\prime})$, we let $(\mathbf{u}_e^{\mathrm{b}\prime}, \widetilde{\boldsymbol{\phi}}_e^{\mathrm{b}\prime}, \boldsymbol{\omega}_e^{\mathrm{b}\prime}) = (\mathbf{u}_e^{\mathrm{r}*}, \widetilde{\boldsymbol{\phi}}_e^{\mathrm{r}*}, \boldsymbol{\omega}_e^{\mathrm{r}*})$ for any $e \in \mathcal{A}^{\mathrm{t}}$ and $e \neq a'$. For $(\mathbf{u}_{a'}^{\mathrm{b}\prime}, \widetilde{\boldsymbol{\phi}}_{a'}^{\mathrm{b}\prime}, \boldsymbol{\omega}_{a'}^{\mathrm{b}\prime})$, we let $u_{a',h'}^{\mathrm{b}\prime} = 1$ and $u_{a',h}^{\mathrm{b}\prime} = 0$ for any $h \in \{1,2,\ldots,H\}$ and $h \neq h'$, and $\omega_{a',h} = \widetilde{\phi}_{a',h} = (\Lambda^{\mathrm{r}*} - \sum_{m=1}^{h'-1} \bar{k}_m \delta)/\bar{k}_{h'}$ and $\omega_{a',h'} = 0$ for any $h \in \{1,2,\ldots,H\}$ and $h \neq h'$. Based on the construction, we have $(\widetilde{\mathbf{Y}}_1^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_2^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_3^{\mathrm{b}\prime}) \in \mathcal{Y}^{\mathrm{b}}$ and the corresponding objective value is $\Theta''(\mathbf{x}) + (\phi_{a'}^{\mathrm{b}\prime} - \phi_{a'}^{\mathrm{r}*})$ by the formulation of (7). Because $(\widetilde{\mathbf{Y}}_1^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_2^{\mathrm{b}\prime}, \widetilde{\mathbf{Y}}_3^{\mathrm{b}\prime}) \in \mathcal{Y}^{\mathrm{b}}$, we also have

$$\Theta''(\mathbf{x}) + (\phi_{a'}^{\mathrm{b}\prime} - \phi_{a'}^{\mathrm{r}*}) \geq \Theta'(\mathbf{x}). \tag{14}$$

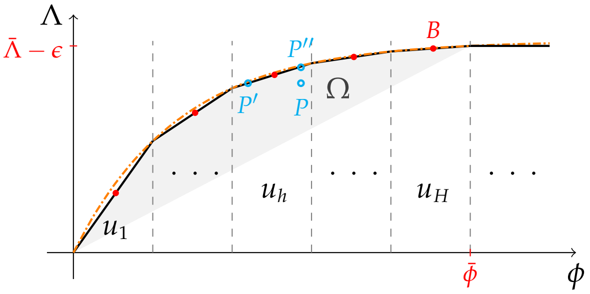


Figure 20 A Geometric Illustration of (ϕ,Λ)

because the objective value with respect to a feasible solution for a minimization problem is no greater than the optimal objective value.

Note that $(\phi_{a'}^{\mathrm{b}\prime},\Lambda_{a'}^{\mathrm{b}\prime})$ amounts to a point in the projection space $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}}}(\mathcal{Y}^{\mathrm{b}})$ and can be illustrated by point P' in Figure 20. For simplicity, we denote the projection space $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{b}}}(\mathcal{Y}^{\mathrm{b}})$, i.e., a piece-wise linear function, by $\Lambda=\Xi(\phi)$, which is also an increasing and concave function. Thus, we have $\Lambda_{a'}^{\mathrm{b}\prime}=\Lambda_{a'}^{r*}=\Xi(\phi_{a'}^{\mathrm{b}\prime})$. Also, there exists $\Lambda^{\mathrm{b}\prime\prime}\in[0,\bar{\Lambda}-\epsilon]$ such that $\Lambda_{a'}^{\mathrm{b}\prime\prime}=\Xi(\phi_{a'}^{r*})$, as illustrated by point P'' in Figure 20. Because the function $\Lambda=\Xi(\phi)$ is concave and $(\phi_{a'}^{\mathrm{r}*},\Lambda_{a'}^{\mathrm{r}*})$ is an interior point of $\mathrm{Proj}_{\widetilde{\mathbf{Y}}_{2,a}^{\mathrm{r}}}(\mathcal{Y}^{\mathrm{r}})$, we have $\Lambda_{a'}^{\mathrm{r}*}<\Lambda_{a'}^{\mathrm{b}\prime\prime}=\Xi(\phi_{a'}^{r*})$. It follows that $\Xi(\phi_{a'}^{\mathrm{b}\prime})<\Xi(\phi_{a'}^{r*})$. Because the function $\Lambda=\Xi(\phi)$ is increasing, we have $\phi_{a'}^{\mathrm{b}\prime}<\phi_{a'}^{r*}$. It further indicates that

$$\Theta''(\mathbf{x})+(\phi_{a'}^{\mathrm{b}\prime}-\phi_{a'}^{\mathrm{r}*})<\Theta''(\mathbf{x}). \tag{15}$$

From (14) and (15), we obtain $\Theta''(\mathbf{x})>\Theta'(\mathbf{x})$, which contradicts with (12). Thus, we have (13) holds and $(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*},\widetilde{\mathbf{Y}}_2^{\mathrm{r}*},\widetilde{\mathbf{Y}}_3^{\mathrm{r}'})\in\mathcal{Y}^{\mathrm{b}}$. It follows that $\Psi(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*},\widetilde{\mathbf{Y}}_2^{\mathrm{r}*},\widetilde{\mathbf{Y}}_3^{\mathrm{r}'})\geq\Theta'(\mathbf{x})$ because of the same reason for (14) to hold. Moreover, by the formulation of (7), where $\widetilde{\mathbf{Y}}_3$ is not involved, we have $\Psi(\widetilde{\mathbf{Y}}_1^{\mathrm{r}*},\widetilde{\mathbf{Y}}_2^{\mathrm{r}*},\widetilde{\mathbf{Y}}_3^{\mathrm{r}'})=\Psi(\widetilde{\mathbf{Y}}^{\mathrm{r}*})=\Theta''(\mathbf{x})$. Thus, we have $\Theta''(\mathbf{x})\geq\Theta'(\mathbf{x})$, which together with (12) imply that $\Theta''(\mathbf{x})=\Theta'(\mathbf{x})$. This completes the proof. □

E.2. Abstract Forms

By (1a), we define $\phi_a^k(\Lambda_a^k)=(-1/\beta_{ij})\ln(1-\Lambda_a^k/\bar{\Lambda}_{ij})$ for any $a=(n_{it},n_{j,t+l_{ij}})\in\mathcal{A}^{\mathrm{t}}$ and $k\in\mathcal{K}$. Since $g(\cdot)$ is bijective, continuous, and concave, we have $\phi_a^k(\cdot)$ is bijective, continuous, and convex. We can then replace ϕ_a^k in (1b) and ($\mathcal{P}$) with $\phi_a^k(\Lambda_a^k)$ and remove constraints (1a), leading to the following abstract form of ($\mathcal{M}$):

$$\Gamma=\min\ \mathbf{c}^\top\mathbf{x}+\sum_{k\in\mathcal{K}}p^k\Big(f(\mathbf{v}^k)+q(\mathbf{\Lambda}^k)+\sum_{a\in\mathcal{A}^{\mathrm{t}}}\phi_a^k(\Lambda_a^k)+\psi(\mathbf{z}^k)\Big) \tag{$\mathcal{M}$}$$

$$\text{s.t. } \mathbf{F}\mathbf{x}\leq\mathbf{S}, \tag{16a}$$

$$\sum_{a\in\mathcal{A}^{\mathrm{t}}}\phi_a^k(\Lambda_a^k)\leq B_{\mathrm{c}},\ k\in\mathcal{K}, \tag{16b}$$

$$\mathbf{A}\mathbf{v}^k+\mathbf{D}\mathbf{\Lambda}^k=\mathbf{C}^k,\ k\in\mathcal{K}, \tag{16c}$$

$$\mathbf{G}\mathbf{z}^k+\mathbf{P}\tilde{\mathbf{z}}^k+\mathbf{L}\mathbf{v}^k\leq\mathbf{Q},\ k\in\mathcal{K}, \tag{16d}$$

$$\mathbf{M}\mathbf{v}^k+\mathbf{U}\mathbf{x}=\mathbf{0},\ k\in\mathcal{K}, \tag{16e}$$

where $f(\cdot)$, $q(\cdot)$, and $\psi(\cdot)$ are linear functions, (16a) represents the first-stage constraints in ($\mathcal{M}$), (16b)–(16c) represent constraints (1b)–(1c) in ($\mathcal{M}$), (16d) represents constraints (2a)–(2g), and (16e) represents constraints (3). We let $\boldsymbol{\rho}^*=(\mathbf{x}^*,\tilde{\mathbf{z}}^*,\mathbf{z}^*,\mathbf{v}^*,\mathbf{\Lambda}^*)$ denote the optimal solution of ($\mathcal{M}$).

For any $a\in\mathcal{A}^{\mathrm{t}}$, $k\in\mathcal{K}$, let $\phi_a^{H,k}(\cdot)$ denote the piecewise-linear (PWL) function with $H>0$ segments that approximates $\phi_a^k(\cdot)$, where $\phi_a^{H,k}(\cdot)$ is continuous and convex and for any $\Lambda_a^k\geq 0$,

$$\phi_a^{H,k}(\Lambda_a^k)\geq\phi_a^k(\Lambda_a^k),\ \lim_{H\to\infty}\phi_a^{H,k}(\Lambda_a^k)=\phi_a^k(\Lambda_a^k). \tag{17}$$

Thus, by replacing $\phi_a^k(\cdot)$ with $\phi_a^{H,k}(\cdot)$ for any $a\in\mathcal{A}^{\mathrm{t}}$, $k\in\mathcal{K}$ in (16b), we have the following abstract form of our PWL approximation model:

$$\Gamma^H=\min\ \mathbf{c}^\top\mathbf{x}+\sum_{k\in\mathcal{K}}p^k\Big(f(\mathbf{v}^k)+q(\mathbf{\Lambda}^k)+\sum_{a\in\mathcal{A}^{\mathrm{t}}}\phi_a^{H,k}(\Lambda_a^k)+\psi(\mathbf{z}^k)\Big) \tag{$\mathcal{M}_H$}$$

$$\text{s.t. } (16a), (16c) - (16e),$$
$$\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^k) \leq B_{\mathrm{c}},\ k \in \mathcal{K}. \tag{18}$$

E.3. Proof of Proposition 3

Proof. Given any $a \in \mathcal{A}^{\mathrm{t}}$, $k \in \mathcal{K}$, and $\epsilon_{1,a}^k > 0$, since $\phi_a^k(\cdot)$ is continuous, there exists $\hat{\delta}_a^k \in (0, \|\Lambda_a^{k^\dagger} - \Lambda_a^{k^*}\|)$ such that $|\phi_a^k(\Lambda_a^k) - \phi_a^k(\Lambda_a^{k^*})| < \epsilon_{1,a}^k$ whenever $\|\Lambda_a^k - \Lambda_a^{k^*}\| \leq \hat{\delta}_a^k$. By the continuity of $q(\cdot)$ and $f(\cdot)$, given any $k \in \mathcal{K}$, $\epsilon_2^k > 0$, and $\epsilon_3^k > 0$, there also exist $\hat{\sigma}^k \in (0, \|\mathbf{\Lambda}^{k^\dagger} - \mathbf{\Lambda}^{k^*}\|)$ and $\hat{\theta}^k \in (0, \|\mathbf{v}^{k^\dagger} - \mathbf{v}^{k^*}\|)$ such that $|q(\mathbf{\Lambda}^k) - q(\mathbf{\Lambda}^{k^*})| < \epsilon_2^k$ and $|f(\mathbf{v}^k) - f(\mathbf{v}^{k^*})| < \epsilon_3^k$ whenever $\|\mathbf{\Lambda}^k - \mathbf{\Lambda}^{k^*}\| < \hat{\sigma}^k$ and $\|\mathbf{v}^k - \mathbf{v}^{k^*}\| < \hat{\theta}^k$, respectively.

First, we construct a feasible solution $\hat{\boldsymbol{\rho}} = (\hat{\mathbf{x}},\ \hat{\tilde{\mathbf{z}}},\ \hat{\mathbf{z}},\ \hat{\mathbf{v}},\ \hat{\mathbf{\Lambda}})$ to model $(\mathcal{M}_H)$. For any $k \in \mathcal{K}$, we define $\hat{\alpha}^k := \min\{\hat{\sigma}^k, \hat{\theta}^k, \hat{\delta}_a^k,\ a \in \mathcal{A}^{\mathrm{t}}\}$ and $\tau^k := (\|\mathbf{\Lambda}^{k^\dagger} - \mathbf{\Lambda}^{k^*}\|^2 + \|\mathbf{v}^{k^\dagger} - \mathbf{v}^{k^*}\|^2)^{\frac{1}{2}}$, and let $\hat{\mathbf{\Lambda}}^k = (\hat{\alpha}^k/\tau^k)\mathbf{\Lambda}^{k^\dagger} + (1 - (\hat{\alpha}^k/\tau^k))\mathbf{\Lambda}^{k^*}$, and $\hat{\mathbf{v}}^k = (\hat{\alpha}^k/\tau^k)\mathbf{v}^{k^\dagger} + (1 - (\hat{\alpha}^k/\tau^k))\mathbf{v}^{k^*}$. Clearly, $\hat{\mathbf{\Lambda}}^k$ (resp. $\hat{\mathbf{v}}^k$) is a convex combination of $\mathbf{\Lambda}^{k^\dagger}$ and $\mathbf{\Lambda}^{k^*}$ (resp. $\mathbf{v}^{k^\dagger}$ and $\mathbf{v}^{k^*}$) for any $k \in \mathcal{K}$. It follows that for any $k \in \mathcal{K}$, $\|\hat{\mathbf{\Lambda}}^k - \mathbf{\Lambda}^{k^*}\| = \hat{\alpha}^k \|\mathbf{\Lambda}^{k^\dagger} - \mathbf{\Lambda}^{k^*}\|/\tau^k \leq \hat{\alpha}^k$ and $\|\hat{\mathbf{v}}^k - \mathbf{v}^{k^*}\| = \hat{\alpha}^k \|\mathbf{v}^{k^\dagger} - \mathbf{v}^{k^*}\|/\tau^k \leq \hat{\alpha}^k$. By the continuity of $\phi_a^k(\cdot)$, $q(\cdot)$, and $f(\cdot)$, we have

$$\left|\phi_a^k(\hat{\Lambda}_a^k) - \phi_a^k(\Lambda_a^{k^*})\right| < \epsilon_{1,a}^k,\ a \in \mathcal{A}^{\mathrm{t}},\ \left|q(\hat{\mathbf{\Lambda}}^k) - q(\mathbf{\Lambda}^{k^*})\right| < \epsilon_2^k,\ \left|f(\hat{\mathbf{v}}^k) - f(\mathbf{v}^{k^*})\right| < \epsilon_3^k,\ k \in \mathcal{K}. \tag{19}$$

We also let $\hat{\mathbf{x}} = \mathbf{x}^\dagger = \mathbf{x}^*$ and $\hat{\tilde{\mathbf{z}}} = \tilde{\mathbf{z}}^\dagger = \tilde{\mathbf{z}}^*$. Note that, by (2b)–(2d) and (2g), $\tilde{\mathbf{z}}$ uniquely determines $\mathbf{z}$. Thus, $\mathbf{z}^\dagger = \mathbf{z}^*$, and we let $\hat{\mathbf{z}} = \mathbf{z}^\dagger = \mathbf{z}^*$. It is easy to check that $\hat{\boldsymbol{\rho}}$ satisfies constraints (16a) and (16c)–(16e) in $(\mathcal{M}_H)$.

In addition, by convexity of $\phi_a^k(\cdot)$, we have $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\hat{\Lambda}_a^k) \leq \sum_{a \in \mathcal{A}^{\mathrm{t}}}[(\hat{\alpha}^k/\tau^k)\phi_a^k(\Lambda_a^{k^\dagger}) + (1 - (\hat{\alpha}^k/\tau^k))\phi_a^k(\Lambda_a^{k^*})] = (\hat{\alpha}^k/\tau^k)\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^\dagger}) + (1 - (\hat{\alpha}^k/\tau^k))\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^*}) < B_{\mathrm{c}}$ for any $k \in \mathcal{K}$, where the last inequality holds because $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^\dagger}) < B_{\mathrm{c}}$ and $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^*}) \leq B_{\mathrm{c}}$. Meanwhile, by (17), there exists $\tilde{H}$ such that $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\hat{\Lambda}_a^k) \leq \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\hat{\Lambda}_a^k) \leq B_c$ for any $H > \tilde{H}$. Thus, $\hat{\boldsymbol{\rho}}$ also satisfies constraints (18) in $(\mathcal{M}_H)$ and $\hat{\boldsymbol{\rho}}$ is feasible to model $(\mathcal{M}_H)$ for any $H > \tilde{H}$.

Second, we prove $\limsup_{H \to \infty} \Gamma^H \leq \Gamma$ in the following. We let $\boldsymbol{\rho}^{H*} = (\mathbf{x}^{H*}, \tilde{\mathbf{z}}^{H*}, \mathbf{z}^{H*}, \mathbf{v}^{H*}, \mathbf{\Lambda}^{H*})$ denote the optimal solution of $(\mathcal{M}_H)$ for any $H > \tilde{H}$. Since $\hat{\boldsymbol{\rho}}$ is feasible to model $(\mathcal{M}_H)$, we have

$$\Gamma^H \leq \mathbf{c}^\top \hat{\mathbf{x}} + \sum_{k \in \mathcal{K}} p^k \Big(f(\hat{\mathbf{v}}^k) + q(\hat{\mathbf{\Lambda}}^k) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\hat{\Lambda}_a^k) + \psi(\hat{\mathbf{z}}^k) \Big). \tag{20}$$

By the proposition condition, for any $H > \max\{\hat{H}, \tilde{H}\}$, we have $\mathbf{x}^{H*} = \mathbf{x}^* = \hat{\mathbf{x}}$, $\tilde{\mathbf{z}}^{H*} = \tilde{\mathbf{z}}^* = \hat{\tilde{\mathbf{z}}}$, and $\mathbf{z}^{H*} = \mathbf{z}^* = \hat{\mathbf{z}}$. Thus, for any $H > \max\{\hat{H}, \tilde{H}\}$, (20) becomes

$$\sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{H,k^*}) + q(\mathbf{\Lambda}^{H,k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{H,k^*}) \Big) \leq \sum_{k \in \mathcal{K}} p^k \Big(f(\hat{\mathbf{v}}^k) + q(\hat{\mathbf{\Lambda}}^k) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\hat{\Lambda}_a^k) \Big). \tag{21}$$

Now, for any $H > \max\{\hat{H}, \tilde{H}\}$, we have

$$\begin{aligned}
\Gamma^H - \Gamma &= \sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{H,k^*}) + q(\mathbf{\Lambda}^{H,k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{H,k^*}) \Big) - \sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{k^*}) + q(\mathbf{\Lambda}^{k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^*}) \Big) \\
&\leq \sum_{k \in \mathcal{K}} p^k \Big(f(\hat{\mathbf{v}}^k) + q(\hat{\mathbf{\Lambda}}^k) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\hat{\Lambda}_a^k) \Big) - \sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{k^*}) + q(\mathbf{\Lambda}^{k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^*}) \Big) \\
&\leq \sum_{k \in \mathcal{K}} p^k \Big(f(\hat{\mathbf{v}}^k) + q(\hat{\mathbf{\Lambda}}^k) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\hat{\Lambda}_a^k) \Big) - \sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{k^*}) + q(\mathbf{\Lambda}^{k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{k^*}) \Big) + \epsilon_0 \\
&\leq \sum_{k \in \mathcal{K}} p^k \Big(\sum_{a \in \mathcal{A}^{\mathrm{t}}} \epsilon_{1,a}^k + \epsilon_2^k + \epsilon_3^k \Big) + \epsilon_0,
\end{aligned}$$

where $\epsilon_0 > 0$ is an arbitrary small number, the first inequality holds by (21), the second inequality holds by (17), and the last inequality holds by (19). Thus, $\limsup_{H \to \infty} \Gamma^H \leq \Gamma$.

Third, for any $H > \max\{\hat{H}, \tilde{H}\}$, $\boldsymbol{\rho}^{H*}$ satisfies constraints (16a) and (16c)–(16e) in $(\mathcal{M})$, and $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{H,k^*}) \leq \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{H,k^*}) \leq B_{\mathrm{c}}$ for any $k \in \mathcal{K}$ by (17). It follows that $\boldsymbol{\rho}^{H*}$ is feasible to model $(\mathcal{M})$ for any $H > \max\{\hat{H}, \tilde{H}\}$. Thus, we have

$$\Gamma \leq \mathbf{c}^\top \mathbf{x}^{H^*} + \sum_{k \in \mathcal{K}} p^k \Big(f(\mathbf{v}^{H,k^*}) + q(\mathbf{\Lambda}^{H,k^*}) + \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{H,k^*}) + \psi(\mathbf{z}^{H,k^*}) \Big). \tag{22}$$

As $\sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{H,k^*}) \leq \sum_{a \in \mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{H,k^*})$ for any $H > \max\{\hat{H}, \tilde{H}\}$, (22) further leads to $\Gamma \leq \Gamma^H$, which, together with $\limsup_{H \to \infty} \Gamma^H \leq \Gamma$, guarantees $\lim_{H \to \infty} \Gamma^H = \Gamma$. □

The condition of Proposition 3 easily holds in general. To show this, we compare the inner approximation $(\mathcal{M}_H)$ with a PWL outer approximation of $(\mathcal{M})$, denoted by $(\mathcal{M}_H^{\mathrm{O}})$. We find that when $H = 20$, the optimality gap between $(\mathcal{M}_H)$ and $(\mathcal{M}_H^{\mathrm{O}})$ is 0.016% and their optimal integer solution of $(\mathbf{x}, \tilde{\mathbf{z}})$ is the same; when H is further increased, this integer solution does not change.

E.4. An Alternative Asymptotic Convergence Result

PROPOSITION 4. *If (i) $B_{\mathrm{c}} = +\infty$ or (ii) $\mathbf{y}^{k^*} > \mathbf{0}$ and $\tilde{\mathbf{x}}^{k^*} > \mathbf{0}$ for any $k \in \mathcal{K}$, then* $\lim_{H\to\infty} \Gamma^H = \Gamma$.

Proof. For the case $B_{\mathrm{c}} = +\infty$, constraints (16b) and (18) can be removed from models ($\mathcal{M}$) and ($\mathcal{M}_H$), which both then share the same set of constraints. We let $\boldsymbol{\rho}^{H*} = (\mathbf{x}^{H*}, \tilde{\mathbf{z}}^{H*}, \mathbf{z}^{H*}, \mathbf{v}^{H*}, \boldsymbol{\Lambda}^{H*})$ denote the optimal solution of ($\mathcal{M}_H$). By (17), we have $\Gamma^H \geq \mathbf{c}^\top \mathbf{x}^{H*} + \sum_{k\in\mathcal{K}} p^k(f(\mathbf{v}^{H,k^*}) + q(\boldsymbol{\Lambda}^{H,k^*}) + \sum_{a\in\mathcal{A}^{\mathrm{t}}} \phi_a^k(\Lambda_a^{H,k^*}) + \psi(\mathbf{z}^{H,k^*})) \geq \Gamma$, where the second inequality holds because $\boldsymbol{\rho}^{H*}$ is feasible to ($\mathcal{M}$).

In addition, we have $\Gamma^H - \Gamma \leq \mathbf{c}^\top \mathbf{x}^* + \sum_{k\in\mathcal{K}} p^k(f(\mathbf{v}^{k^*}) + q(\boldsymbol{\Lambda}^{k^*}) + \sum_{a\in\mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{k^*}) + \psi(\mathbf{z}^{k^*})) - \Gamma = \sum_{k\in\mathcal{K}} p^k(\sum_{a\in\mathcal{A}^{\mathrm{t}}}(\phi_a^{H,k}(\Lambda_a^{k^*}) - \phi_a^k(\Lambda_a^{k^*})))$, where the inequality holds because $\boldsymbol{\rho}^*$ is feasible to model ($\mathcal{M}_H$). By (17), for any $\epsilon_H^k > 0$, there exists $\tilde{H}$ such that $|\phi_a^{H,k}(\Lambda_a^{k^*}) - \phi_a^k(\Lambda_a^{k^*})| < \epsilon_H^k$ for any $H > \tilde{H}$. Thus, for any $H > \tilde{H}$, $\Gamma^H - \Gamma \leq \sum_{k\in\mathcal{K}} p^k |\mathcal{A}^{\mathrm{t}}| \epsilon_H^k$, which, together with $\Gamma^H \geq \Gamma$, guarantees $\lim_{H\to\infty} \Gamma^H = \Gamma$ because $\mathcal{K}$ is finite and p^k, $k \in \mathcal{K}$ and $|\mathcal{A}^{\mathrm{t}}|$ are bounded.

For the case $\mathbf{y}^{k^*} > \mathbf{0}$ and $\tilde{\mathbf{x}}^{k^*} > \mathbf{0}$ for any $k \in \mathcal{K}$, given $H > \tilde{H}$ and based on the optimal solution $\boldsymbol{\rho}^*$ to model ($\mathcal{M}$), we construct a feasible solution $\boldsymbol{\rho}^H = (\mathbf{x}^H, \tilde{\mathbf{z}}^H, \mathbf{z}^H, \mathbf{v}^H, \boldsymbol{\Lambda}^H)$ to model ($\mathcal{M}_H$) such that $\mathbf{x}^H = \mathbf{x}^*$, $\tilde{\mathbf{z}}^H = \tilde{\mathbf{z}}^*$, $\mathbf{z}^H = \mathbf{z}^*$, and

$$\phi_a^{H,k}(\Lambda_a^{H,k}) = \phi_a^k(\Lambda_a^{k^*}),\ a \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}. \tag{23}$$

By (17), we have $\Lambda_a^{k^*} \geq \Lambda_a^{H,k}$, $a \in \mathcal{A}^{\mathrm{t}}$, $k \in \mathcal{K}$, by which we define $\bar{\delta}_a^{H,k} := \Lambda_a^{k^*} - \Lambda_a^{H,k} \geq 0$. If $\bar{\delta}_a^{H,k} > 0$ for some $a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$ and $k \in \mathcal{K}$, then by the mean value theorem, there exists $\Lambda \in (\Lambda_a^{H,k}, \Lambda_a^{k^*})$ such that the derivative $\phi_a^{k'}(\Lambda) > 0$ and $\phi_a^{k'}(\Lambda)(\Lambda_a^{k^*} - \Lambda_a^{H,k}) = \phi_a^k(\Lambda_a^{k^*}) - \phi_a^k(\Lambda_a^{H,k})$. Therefore,

$$\begin{aligned}\Lambda_a^{k^*} - \Lambda_a^{H,k} &= \frac{1}{\phi_a^{k'}(\Lambda)}\left(\phi_a^k(\Lambda_a^{k^*}) - \phi_a^k(\Lambda_a^{H,k})\right) = \beta_{ij}(\bar{\Lambda}_{ij} - \Lambda)\left(\phi_a^k(\Lambda_a^{k^*}) - \phi_a^k(\Lambda_a^{H,k})\right) \\ &= \beta_{ij}(\bar{\Lambda}_{ij} - \Lambda)\left(\phi_a^{H,k}(\Lambda_a^{H,k}) - \phi_a^k(\Lambda_a^{H,k})\right),\end{aligned} \tag{24}$$

where the last equality holds by (23). By (24) and (17), we further have

$$\bar{\delta}_a^{H,k} = \Lambda_a^{k^*} - \Lambda_a^{H,k} \leq \beta_{ij}\bar{\Lambda}_{ij}\epsilon_H^k := \bar{\epsilon}_H^k,\ a = (n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}, \tag{25}$$

for any $H > \tilde{H}$, i.e., $\lim_{H\to\infty} \bar{\delta}_a^{H,k} = 0$ because $\beta_{ij},\ \bar{\Lambda}_{ij},\ j \neq i,\ i, j \in \mathcal{V}$ are bounded.

We further construct $\mathbf{v}^H = (\widetilde{\mathbf{x}}^H, \boldsymbol{\eta}^H, \mathbf{y}^H, \boldsymbol{\gamma}^H)$ as follows. First, we let $\boldsymbol{\gamma}^H = \boldsymbol{\gamma}^*$. Second, we let $\eta_a^{H,k} = \eta_a^{k^*} - \zeta_a^{H,k},\ a \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}$, where $\zeta_a^{H,k} = \bar{\delta}_a^{H,k}$ if $\eta_a^{k^*} > 0$ and $\zeta_a^{H,k} = 0$ otherwise. Here we define $\theta_a^{H,k} = \bar{\delta}_a^{H,k} - \zeta_a^{H,k},\ a \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}$. Third, for any $a \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}$, we let $y_a^{H,k} = y_a^{k^*} - \theta_a^{H,k}$, which is nonnegative by (1c). Fourth, we let

$$y_{n_{it},n_{i,t+1}}^{H,k} := y_{n_{it},n_{i,t+1}}^{k^*} + \sum_{s=0}^{t} \sum_{j\in\mathcal{V},j\neq i} \theta_{n_{is},n_{j,s+l_{ij}}}^{H,k} - \sum_{s=1}^{t} \sum_{j\in\mathcal{V},j\neq i} \theta_{n_{j,s-l_{ji}},n_{is}}^{H,k},\ i \in \mathcal{V},\ t \in \{0, \cdots, T-2\},\ k \in \mathcal{K}.$$

Since T and $|\mathcal{V}|$ are bounded, by the propsition condition and $\theta_a^{H,k} \leq \bar{\delta}_a^{H,k} \leq \bar{\epsilon}_H^k,\ a \in \mathcal{A}^{\mathrm{t}},\ k \in \mathcal{K}$, we have $y_{n_{it},n_{i,t+1}}^{H,k} \geq 0,\ i \in \mathcal{V},\ t \in \{0, \cdots, T-2\},\ k \in \mathcal{K}$. Fifth, we let

$$\tilde{x}_i^{H,k} = \tilde{x}_i^{k^*} + \sum_{s=0}^{T-2} \sum_{j\in\mathcal{V},j\neq i} \theta_{n_{is},n_{j,s+l_{ij}}}^{H,k} - \sum_{s=1}^{T-1} \sum_{j\in\mathcal{V},j\neq i} \theta_{n_{j,s-l_{ji}},n_{is}}^{H,k},\ \ i \in \mathcal{V},\ k \in \mathcal{K},$$

where we have $\tilde{x}_i^{H,k} \geq 0$ similarly.

It is easy to check that $\boldsymbol{\rho}^H$ is feasible to model ($\mathcal{M}_H$). Thus, for any $H > \tilde{H}$, we have

$$\begin{aligned}0 &\leq \Gamma^H - \Gamma \leq \mathbf{c}^\top \mathbf{x}^H + \sum_{k\in\mathcal{K}} p^k(f(\mathbf{v}^{H,k}) + q(\boldsymbol{\Lambda}^{H,k}) + \sum_{a\in\mathcal{A}^{\mathrm{t}}} \phi_a^{H,k}(\Lambda_a^{H,k}) + \psi(\mathbf{z}^{H,k})) - \Gamma \\ &= \sum_{k\in\mathcal{K}} p^k \left(\sum_{a\in\mathcal{A}^{\mathrm{t}}} \left(C_{\mathrm{p}}(\eta_a^{H,k} - \eta_a^{k^*}) - Rl_a(y_a^{H,k} - y_a^{k^*}) + Rl_a(\Lambda_a^{H,k} - \Lambda_a^{k^*}) \right) \right) \leq \sum_{k\in\mathcal{K}} p^k \Big(\sum_{a\in\mathcal{A}^{\mathrm{t}}} Rl_a(y_a^{k^*} - y_a^{H,k}) \Big) = \sum_{k\in\mathcal{K}} p^k \Big(\sum_{a\in\mathcal{A}^{\mathrm{t}}} Rl_a \theta_a^{H,k} \Big) \\ &\leq \sum_{k\in\mathcal{K}} p^k \Big(\sum_{a\in\mathcal{A}^{\mathrm{t}}} Rl_a \bar{\delta}_a^{H,k} \Big) \leq \sum_{k\in\mathcal{K}} p^k \Big(\sum_{i\in\mathcal{V}} \sum_{j\neq i,\, j\in\mathcal{V}} \sum_{t\in\mathcal{T}(l_{ij})} Rl_{ij}\beta_{ij}\bar{\Lambda}_{ij}\epsilon_H^k \Big),\end{aligned}$$

where the first inequality holds because model ($\mathcal{M}_H$) is an inner approximation of model ($\mathcal{M}$), the second inequality holds because $\boldsymbol{\rho}^H$ is feasible to model ($\mathcal{M}_H$), the third inequality holds because $\eta_a^{H,k} \leq \eta_a^{k^*}$ and $\Lambda_a^{H,k} \leq \Lambda_a^{k^*}$, and the last inequality holds by (25). Thus, $\lim_{H\to\infty} \Gamma^H = \Gamma$. □

E.5. The Refined Model

Given a starting period $s \in \mathcal{T}$ and ending period $e \in \mathcal{T} \cup \{T\}$, we define $\mathcal{A}^{\mathrm{t}}(s,e)$ as the subset of $\mathcal{A}^{\mathrm{t}}$ in period $t \in \mathcal{T}(l_{ij}, s, e)$ and denote the second-stage variables between the time range $[s,e)$ in each scenario $k \in \mathcal{K}$ by vector $\widetilde{\mathbf{Y}}^k_{(s,e)} = (\widetilde{\mathbf{x}}^k, \boldsymbol{\eta}^k, \mathbf{y}^k, \boldsymbol{\gamma}^k, \boldsymbol{\Lambda}^k, \mathbf{u}^k, \boldsymbol{\phi}^k, \widetilde{\boldsymbol{\phi}}^k, \mathbf{z}^k, \widetilde{\mathbf{z}}^k, \boldsymbol{\omega}^k)_{(s,e)}$. Then, we can write $\mathcal{Q}(s,e)$ as follows:

$$\begin{aligned} \Theta''_{(s,e)}(\mathbf{x}) = \min_{\mathbf{Y}^k_{(s,e)},\, k \in \mathcal{K}} \quad & \sum_{k \in \mathcal{K}} p^k \Bigg(\sum_{a \in \mathcal{A}^{\mathrm{t}}(s,e)} \Big(\mathrm{C_p} \eta^k_a + \phi^k_a - Rl_a (y^k_a - \Lambda^k_a) \Big) + \sum_{t \in \mathcal{T}(l_{\mathrm{r}}, s, e)} \mathrm{C_r} z^k_t \Bigg) \\ \text{s.t.} \quad & \mathbf{Y}^k_{(s,e)} \in \mathcal{Y}_{(s,e)}(\mathbf{x}, \boldsymbol{\lambda}^k),\ k \in \mathcal{K}. \end{aligned} \tag{$\mathcal{Q}(s,e)$}$$

For ease of notation, we omit superscript k for variables and parameters in $\mathcal{Y}_{(s,e)}(\mathbf{x}, \boldsymbol{\lambda}^k)$ for any $k \in \mathcal{K}$. Then, the feasible region $\mathcal{Y}_{(s,e)}(\mathbf{x}, \boldsymbol{\lambda})$ with respect to each scenario is given by

$$\mathcal{Y}_{(s,e)}(\mathbf{x}, \boldsymbol{\lambda}) := \quad \Big\{ \widetilde{\mathbf{Y}}_{(s,e)} \geq \mathbf{0} \,\Big|\, \Big(y_{n_{it}, n_{i,t+1}} + \sum_{j \in \mathcal{V}, j \neq i} \Big(y_{n_{it}, n_{j,t+l_{ij}}} + \gamma_{n_{it}, n_{j,t+l_{\mathrm{r}}}} \Big) \Big) - \Big(y_{n_{i,t-1}, n_{i,t}} \tag{26a}$$

$$+ \sum_{j \in \mathcal{V}, j \neq i} \Big(y_{n_{j,t-l_{ij}}, n_{i,t}} + \gamma_{n_{j,t-l_{\mathrm{r}}}, n_{i,t}} \Big) \Big) = \begin{cases} x_i, & \text{if } t = s, \\ 0, & \text{if } t = s+1, \ldots, e-2, \quad i \in \mathcal{V}, \\ -\widetilde{x}_i, & \text{if } t = e-1, \end{cases} \tag{26b}$$

$$\lambda_{n_{it}, n_{j,t+l_{ij}}} + \Lambda_{n_{it}, n_{j,t+l_{ij}}} - \eta_{n_{it}, n_{j,t+l_{ij}}} = y_{n_{it}, n_{j,t+l_{ij}}},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26c}$$

$$\phi_{n_{it}, n_{j,t+l_{ij}}} = \sum_{h=1}^{H} (h-1) u_{h, n_{it}, n_{j,t+l_{ij}}} \delta_{ij} + \omega_{h, n_{it}, n_{j,t+l_{ij}}},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26d}$$

$$\Lambda_{n_{it}, n_{j,t+l_{ij}}} = \sum_{h=1}^{H} u_{h, n_{it}, n_{j,t+l_{ij}}} \sum_{m=1}^{h-1} \bar{k}_{mij} \delta_{ij} + \sum_{h=1}^{H} \bar{k}_{h, n_{it}, n_{j,t+l_{ij}}} \omega_{h, n_{it}, n_{j,t+l_{ij}}},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26e}$$

$$u_{h, n_{it}, n_{j,t+l_{ij}}} \delta_{ij} - \omega_{h, n_{it}, n_{j,t+l_{ij}}} \geq 0,\ h \in \mathcal{H},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26f}$$

$$\widetilde{\phi}_{h, n_{it}, n_{j,t+l_{ij}}} - \omega_{h, n_{it}, n_{j,t+l_{ij}}} \geq 0,\ h \in \mathcal{H},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26g}$$

$$\delta_{ij} - u_{h, n_{it}, n_{j,t+l_{ij}}} \delta_{ij} - \widetilde{\phi}_{h, n_{it}, n_{j,t+l_{ij}}} + \omega_{h, n_{it}, n_{j,t+l_{ij}}} \geq 0,\ h \in \mathcal{H},\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26h}$$

$$\sum_{h=1}^{H} u_{h, n_{it}, n_{j,t+l_{ij}}} = 1,\ t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26i}$$

$$\widetilde{\phi}_{h, n_{it}, n_{j,t+l_{ij}}} \leq \delta_{ij},\ h \in \mathcal{H}, t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26j}$$

$$u_{h, n_{it}, n_{j,t+l_{ij}}} \leq 1,\ h \in \mathcal{H}, t \in \mathcal{T}(l_{ij}, s, e), j \neq i, i, j \in \mathcal{V}, \tag{26k}$$

$$\sum_{i=t}^{t+l_{\mathrm{f}}-1} \widetilde{z}_i \leq \bar{\bar{z}},\ t \in \{s, \ldots, e - l_{\mathrm{r}} - l_{\mathrm{f}}\}, \tag{26l}$$

$$\sum_{t \in \mathcal{T}(l_{\mathrm{r}}, s, e)} z_t \leq \bar{z}, \tag{26m}$$

$$\underline{q} \widetilde{z}_t \leq \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}, j \neq i} \gamma_{n_{it}, n_{j,t+l_{\mathrm{r}}}} \leq \bar{q} \widetilde{z}_t,\ t \in \mathcal{T}(l_{\mathrm{r}}, s, e), \tag{26n}$$

$$z_t - \widetilde{z}_t \geq 0,\ t = s;\ \ z_t - \widetilde{z}_t + \widetilde{z}_{t-1} \geq 0,\ t \in \mathcal{T}(l_{\mathrm{r}}, s+1, e), \tag{26o}$$

$$z_t + \widetilde{z}_{t-1} - 1 \leq 0,\ z_t - \widetilde{z}_t \leq 0,\ t \in \mathcal{T}(l_{\mathrm{r}}, s, e), \tag{26p}$$

$$\widetilde{z}_t \in \{0, 1\},\ t \in \mathcal{T}(l_{\mathrm{r}}, s, e) \Big\}. \tag{26q}$$

Note that $\widetilde{\mathbf{Y}}_{(s,e)} \geq \mathbf{0}$ indicates every variable in the vector $\widetilde{\mathbf{Y}}_{(s,e)}$ is non-negative. Finally, the two-stage stochastic program $\mathcal{M}(s,e)$ is given by

$$\min_{\mathbf{x}} \sum_{j \in \mathcal{V}} c_j x_j + \Theta''(\mathbf{x}) \ \text{ s.t. } \ \mathbf{x} \in \mathcal{X} = \left\{ \mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \;\middle|\; x_j - B_j \leq 0,\ j \in \mathcal{V},\ \sum_{j \in \mathcal{V}} x_j \leq N \right\}. \tag{$\mathcal{M}(s,e)$}$$

E.6. Algorithm 1 Details

For any $s \in \mathcal{T}$ and $e \in \mathcal{T} \cup \{T\}$ such that $e - s \geq 3$, we fix some binary variables in problem $\mathcal{M}(s,e)$ to be 0 based on its LP relaxation solution. In Algorithm 1, we apply the following procedure for each scenario $k \in \mathcal{K}$. First, we construct two sets (i.e., $\mathcal{T}^{\mathrm{r}1}$ and $\mathcal{T}^{\mathrm{r}2}$) by, respectively, collecting the following two sets of values: (i) $\tilde{z}_t^{ks},\ t \in \mathcal{T}(l_{\rm r},s,e)$ and (ii) $\hat{z}_t^{ks} = \tilde{z}_t^{ks} + \tilde{z}_{t+1}^{ks},\ t \in \mathcal{T}(l_{\rm r},s,e-1)$. Each of these two sets of values are sorted in decreasing order in sets $\mathcal{T}^{\rm r1}$ and $\mathcal{T}^{\rm r2}$. Next, we collect the time indices (i.e., t) of the first $n_I \le e-s-1$ elements of the ordered set $\mathcal{T}^{\rm r1}$ in $\hat{\mathcal{T}}^{\rm r1}$ and those of $\mathcal{T}^{\rm r2}$ in $\hat{\mathcal{T}}^{\rm r2}$. By increasing the value of each element in $\hat{\mathcal{T}}^{\rm r2}$ by 1, we collect these indices in $\hat{\mathcal{T}}^{\rm r3}$. Note that for each scenario $k \in \mathcal{K}$, the subset of variables $\{\tilde{z}_t^k,\ t \in \hat{\mathcal{T}}^{\rm r1} \cup \hat{\mathcal{T}}^{\rm r2} \cup \hat{\mathcal{T}}^{\rm r3}\}$ are most likely to take value 1. Thus, we reduce the binary variables $\tilde{z}_t^k$ in the operational problem $\mathcal{Q}_{\rm LP}(s,e)$ and $\mathcal{Q}(s,e)$ by restricting $\tilde{z}_t^k = 0$ if $t \in \mathcal{T}(l_{\rm r},s,e) \backslash (\hat{\mathcal{T}}^{\rm r1} \cup \hat{\mathcal{T}}^{\rm r2} \cup \hat{\mathcal{T}}^{\rm r3})$ for any scenario $k \in \mathcal{K}$. By (2f), when $\tilde{z}_t^k = 0$, the number of vehicles relocated by the 3PL $\gamma_{n_{it},n_{j,t+l_{\rm r}}} = 0$ from region i in period t to region j. Thus, the solution space is reduced and the solution process can be accelerated.

Algorithm 1 Reduction of the Binary Variables for 3PL Relocation (s,e)

```
Initialize: $\mathcal{T}^{\rm r1} = \hat{\mathcal{T}}^{\rm r1} = \mathcal{T}^{\rm r2} = \hat{\mathcal{T}}^{\rm r2} = \hat{\mathcal{T}}^{\rm r3} := \emptyset$
Solve $\mathcal{M}_{\rm LP}(s,e)$ to obtain $\tilde{z}_t^{ks},\ t \in \mathcal{T}(l_{\rm r},s,e),\ k \in \mathcal{K}$
for $k \in \mathcal{K}$ do
    $\mathcal{T}^{\rm r1} \leftarrow \{\tilde{z}_t^{ks} \mid t \in \mathcal{T}(l_{\rm r},s,e)\}$ such that $\tilde{z}_{(1)}^{ks} \ge \cdots \ge \tilde{z}_{(e-s-l_{\rm r}-1)}^{ks}$ and $\mathcal{T}^{\rm r1} = \{\tilde{z}_{(t)}^{ks} \mid t \in \mathcal{T}(l_{\rm r},1,e-s)\}$
    $\hat{z}_t^{ks} \leftarrow \tilde{z}_t^{ks} + \tilde{z}_{t+1}^{ks},\ t \in \mathcal{T}(l_{\rm r},s,e-1)$
    $\mathcal{T}^{\rm r2} \leftarrow \{\hat{z}_t^{ks} \mid t \in \mathcal{T}(l_{\rm r},s,e)\}$ such that $\hat{z}_{(1)}^{ks} \ge \cdots \ge \hat{z}_{(e-s-l_{\rm r}-2)}^{k*}$ and $\mathcal{T}^{\rm r2} = \{\hat{z}_{(t)}^{ks} \mid t \in \mathcal{T}(l_{\rm r},1,e-s-1)\}$
    $\hat{\mathcal{T}}^{\rm r1} \leftarrow \{t \in \mathcal{T}(l_{\rm r},s,e) \mid \tilde{z}_t^{ks} \ge \tilde{z}_{(n_I)}^{ks}\}$, $\hat{\mathcal{T}}^{\rm r2} \leftarrow \{t \in \mathcal{T}(l_{\rm r},s,e-1) \mid \hat{z}_t^{ks} \ge \hat{z}_{(n_I)}^{ks}\}$, $\hat{\mathcal{T}}^{\rm r3} \leftarrow \{t+1 \mid t \in \hat{\mathcal{T}}^{\rm r2}\}$
    Set $\tilde{z}_t^k = 0,\ t \in \mathcal{T}(l_{\rm r},s,e) \backslash (\hat{\mathcal{T}}^{\rm r1} \cup \hat{\mathcal{T}}^{\rm r2} \cup \hat{\mathcal{T}}^{\rm r3})$ in $\mathcal{Q}_{\rm LP}(s,e)$ and $\mathcal{Q}(s,e)$
    $\mathcal{T}^{\rm r1}, \hat{\mathcal{T}}^{\rm r1}, \mathcal{T}^{\rm r2}, \hat{\mathcal{T}}^{\rm r2}, \hat{\mathcal{T}}^{\rm r3} \leftarrow \emptyset$
```

E.7. Algorithm 2 Details

Algorithm 2 Temporal Decomposition

```
Initialize: $\underline{\mathbf{x}} := \mathbf{0}$, $x_{\rm sum} := 0$, $\mathcal{D}_m := \emptyset,\ m \in \{1,2,\ldots,M\}$
for $m = M,\ldots,1$ do
    $s \leftarrow (m-1)T_{\rm sub}$, $e \leftarrow mT_{\rm sub}$
    Update problem $\mathcal{Q}_{\rm LP}(s,e)$ and $\mathcal{Q}(s,e)$ to reduce binary variables by Algorithm 1
    Solve $\mathcal{M}_{\rm LP}(s,e)$ with $\widetilde{\mathbf{Y}}_{(s,e)} \in \mathcal{Y}_{{\rm LP}(s,e)}(\mathbf{x},\boldsymbol{\lambda}) \cap \{\widetilde{\mathbf{Y}}_{(s,e)} \mid \tilde{x}_i \ge \underline{x}_i,\ i \in \mathcal{V}\}$ to obtain solution $\mathbf{x}^s$ and objective $v^s$
    $\mathcal{D}_m \leftarrow \mathcal{D}_m \cup \{(\mathbf{x}^s, v^s)\}$, $\underline{x}_i \leftarrow x_i^s$, $x_{\rm sum} \leftarrow \sum_{i \in \mathcal{V}} x_i^s$
    for $\sigma_{\rm sum} = \underline{\sigma},\ldots,\bar{\sigma}$ do
        $\mathcal{X}' \leftarrow \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \sum_{i \in \mathcal{V}} x_i = x_{\rm sum} + \sigma_{\rm sum},\ \underline{x}_i - \sigma_i \le x_i \le \underline{x}_i + \sigma_i,\ i \in \mathcal{V}\}$
        Resolve $\mathcal{M}_{\rm LP}(s,e)$ where $\mathbf{x} \in \mathcal{X}'$ by Bender's decomposition
        Obtain solution $\mathbf{x}^s$ and objective $v^s$
        $\mathcal{D}_m \leftarrow \mathcal{D}_m \cup \{(\mathbf{x}^s, v^s)\}$, $\mathcal{X} \leftarrow \emptyset$
    if $m \ne 1$ then
        $\mathcal{D}_m \leftarrow \{(\mathbf{x}_\ell, v_\ell) \in \mathcal{D}_m,\ \ell \in \{1,2,\ldots,|\mathcal{D}_m|\} | v_1 \le \cdots \le v_{|\mathcal{D}_m|}\}$
        $\mathbf{x}^s \leftarrow \mathcal{D}_m[0].key$, $\underline{\mathbf{x}} \leftarrow \mathbf{x}^s$
    else
        for all $(\mathbf{x}^s, v^s) \in \mathcal{D}_1$ do
            $\hat{\mathcal{T}} \leftarrow \{(t,k) \in \mathcal{T}(l_{\rm r},0,T) \times \mathcal{K} \mid \tilde{z}_t^k = 0$ in $\mathcal{Q}(s,e)$ with $t \in \mathcal{T}(l_{\rm r},s,e)$ where $s = (m-1)T_{\rm sub},\ e = mT_{\rm sub},\ m \in \{1,2,\ldots,M\}\}$
            Resolve $\mathcal{M}(0,T)$ where $\mathbf{x} \in \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \mathbf{x} = \mathbf{x}^s\}$ and $\widetilde{\mathbf{Y}} \in \mathcal{Y}(\mathbf{x},\boldsymbol{\lambda}) \cap \{\widetilde{\mathbf{Y}} \mid \tilde{z}_t^k = 0, (t,k) \in \hat{\mathcal{T}}\}$ to obtain $v'$
            $v^s \leftarrow v'$, update $(\mathbf{x}^s, v^s)$ in $\mathcal{D}_1$
Return $\mathcal{D}_1$
```

In Algorithm 2, we solve smaller and thus simpler two-stage stochastic programs over each sub-network $m \in \{1,2,\ldots,M\}$ in a backward sequence, from the last sub-network $m = M$ to the first $m = 1$. For each sub-network, we follow a four-step solution procedure, with the first step to reduce the number of second-stage binary variables by applying Algorithm 1. In the following, we detail the remaining three steps for solving the two-stage stochastic programs (i.e., $\mathcal{M}((m-1)T_{\rm sub}, mT_{\rm sub})$) over the sub-networks $m \ge 2$ and $m = 1$ separately. For any $m \ge 2$, we have the following detailed Steps 2 - 4:

- Step 2. When solving the two-stage problem $\mathcal{M}_{\rm LP}(s,e)$ over the sub-network m (i.e., $s = (m-1)T_{\rm sub}$, $e = mT_{\rm sub}$), we add the constraints $\tilde{x}_i \ge \underline{x}_i,\ i \in \mathcal{V}$, in the second-stage problem, i.e., the outflow $\tilde{x}_i$ of the sub-network m is lower bounded by $\underline{x}_i$. This lower bound takes the value of the first-stage solution $x_i^s,\ i \in \mathcal{V}$, of the two-stage problem $\mathcal{M}_{\rm LP}(s,e)$ over the sub-network $m+1$ if $m < M$ (i.e.,

$s = mT_{\mathrm{sub}}$, $e = (m+1)T_{\mathrm{sub}}$) and takes the value of 0 if $m = M$. The intuition behind such constraints is that (i) the vehicle allocation solution of the two-stage problem over the sub-network m approximates that over the sub-networks from m to M; (ii) the demand over sub-networks from m to M is no less than that over the sub-networks from $m+1$ to M and the higher demand often leads to more allocated vehicles. Then, we derive the first-stage allocation solution to $\mathcal{M}_{\mathrm{LP}}(s,e)$ over the sub-network m.

- Step 3. We refine and re-solve the problem $\mathcal{M}_{\mathrm{LP}}(s,e)$ several more times, where for each time, two types of constraints are added, yielding several candidate allocations. One type of constraints are to perturb the total number of allocated vehicles to take a different value. Particularly, we set the total number of allocated vehicles $\sum_{i\in\mathcal{V}} x_i = x_{\mathrm{sum}} + \sigma_{\mathrm{sum}}$ for each $\sigma_{\mathrm{sum}} \in \{\underline{\sigma}, \ldots, \bar{\sigma}\}$, where σ_{sum} denotes the perturbed value. The others are box constraints $x_i \in [\underline{x}_i - \sigma_i, \underline{x}_i + \sigma_i]$, $i \in \mathcal{V}$, where σ_i (with $\sum_{i\in\mathcal{V}} \sigma_i \leq \sigma_{\mathrm{sum}}$) is the maximum purterbed value for the regional vehicle allocation. Slightly perturbing the total number of allocated vehicles to different potential values has two advantages. First, it can reduce the feasible region of the vehicle allocation variables. Second, we observe that, when the total number of allocated vehicles is fixed, a small perturbation of the total number of vehicles allocated to one or more regions has a minor impact on the objective value of the integrated allocation and relocation problem. The reason is intuitive: though regional vehicle allocation is perturbed, the vehicles can be relocated to proper places at a low cost. However, when the total number of allocated vehicles changes, the objective value of the two-stage problem changes significantly.
- Step 4. With several candidate solutions for allocating vehicles in sub-network m, we store each solution $\mathbf{x}^s$ as the *key* and the objective value v^s as the *value* in dictionary $\mathcal{D}_m$, which stores data by the key-value pairs and the subscript m represents the sub-network index. Then, the solution $\mathbf{x}^s$ with the smallest objective v^s in $\mathcal{D}_m$ is chosen and for each $i \in \mathcal{V}$, x_i^s is taken as the lower bound $\underline{x}_i$ for the two-stage problem over sub-network $m-1$.

When $m = 1$, a similar four-step solution procedure is applied as above except that in the final step, we do not choose the x_i^s with the smallest objective value in $\mathcal{D}_1$. Instead, for each candidate solution $\mathbf{x}^s$ in $\mathcal{D}_1$, a two-stage problem $\mathcal{M}(0,T)$ over the entire network is solved with the first-stage decisions fixed as $\mathbf{x}^s$ and the number of second-stage binary variables reduced. The resulting objective value is denoted by v' and used to update v^s with respect to $\mathbf{x}^s$. We thus obtain several candidate solutions $\mathbf{x}^s$ and the corresponding objectives value v^s in $\mathcal{D}_1$ for the two-stage problem $\mathcal{M}(0,T)$ over the original time-space network.

We note that the network split in Algorithm 2 will break some original arcs connecting the sub-networks. However, there are periods when the shared micromobility system has a relatively low demand and thus a low volume of traffic on the broken arcs. Thus, the temporal decomposition shall cause a minor error when splitting the network in those low-demand periods.

Whenever solving a two-stage problem, we apply Bender's decomposition such that the second-stage problem is further decomposed into independent scenario-based subproblems, which is solved in parallel. Our temporal decomposition algorithm can efficiently solve large-scale problems because (i) the problem size at each iteration is reduced, so is the number of binary variables; (ii) the solution space of $\mathcal{M}_{\mathrm{LP}}(s,e)$ is reduced by fixing $\sum_{i\in\mathcal{V}} x_i$ to different values, thereby enabling a quick search of better solutions; and (iii) by employing Bender's decomposition, we can leverage the parallel capacity to solve the second-stage problem $\mathcal{P}^k$ for each separate scenario $k \in \mathcal{K}$ in parallel, further enhancing the computational efficiency.

Furthermore, we formulate the Bender's cuts used in the above Algorithm 2. We note that our two-stage problem $\mathcal{M}_{\mathrm{LP}}(s,e)$ over any sub-network has a complete recourse, and hence we only need to derive the Bender's optimality cuts. Given the operational horizon $[s,e)$, denote the arc set $\mathcal{A}^{\mathrm{t}}$ with $t \in \mathcal{T}(l_{ij}, s, e)$ by $\mathcal{A}^{\mathrm{t}}(s,e)$. Also, given the first-stage solution $\mathbf{x}$, denote the multipliers for the first case of constraints (26b) (with right-hand-side x_i) by $\pi_i^1 \in \mathbb{R}$, $i \in \mathcal{V}$, those for constraints (26l) by $\pi_t^2 \geq 0$, $t \in \{s, \ldots, e - l_{\mathrm{r}} - l_{\mathrm{f}}\}$, that for constraint (26m) by $\pi^3 \geq 0$, those for constraints (26i) by π_a^4, $a \in \mathcal{A}^{\mathrm{t}}(s,e)$, those for constraints (26j) by $\pi_{h,a}^5 \geq 0$, $h \in \mathcal{H}$, $a \in \mathcal{A}^{\mathrm{t}}(s,e)$, those for constraints (26h) by $\pi_{h,a}^6 \geq 0$, $h \in \mathcal{H}$, $a \in \mathcal{A}^{\mathrm{t}}(s,e)$, those for (26k) by $\pi_{h,a}^7 \geq 0$, $h \in \mathcal{H}$, $a \in \mathcal{A}^{\mathrm{t}}(s,e)$, those for constraints (26c) by $\pi_a^8 \in \mathbb{R}$, $a \in \mathcal{A}^{\mathrm{t}}(s,e)$, those for the first part of constraints (26p) ($z_t + \widetilde{z}_{t-1} - 1 \leq 0$) by $\pi_t^9 \geq 0$, $t \in \mathcal{T}(l_{\mathrm{r}}, s+1, e)$, those for $\tilde{z}_t \leq 1$ by $\pi_t^{10} \geq 0$, $t \in \mathcal{T}(l_{\mathrm{r}}, s, e)$, and those for constraints $\tilde{x}_i \geq \underline{x}_i$ (added in Algorithm 2) by $\pi_i^{11} \geq 0$, $i \in \mathcal{V}$. We omit the dual variables for other constraints because they do not appear in the optimality cuts. At a given iteration, a Bender's optimality cut takes the form $\theta - \sum_{k\in\mathcal{K}} p^k q^k(\mathbf{x}) \geq 0$, where θ denotes the lower bound approximation of the second-stage value function $\Theta''_{(s,e)}(\mathbf{x})$ and

$$q^k(\mathbf{x}) = \sum_{i\in\mathcal{V}} \pi_i^1 x_i - \sum_{t=s}^{e-l_{\mathrm{r}}-l_{\mathrm{f}}} \pi_t^2 \overset{\equiv}{z} - \pi^3 \bar{z} + \sum_{a\in\mathcal{A}^{\mathrm{t}}(s,e)} \pi_a^4 - \sum_{h\in\mathcal{H}, a\in\mathcal{A}^{\mathrm{t}}(s,e)} \left(\left(\pi_{h,a}^5 + \pi_{h,a}^6 \right) \delta_{ij} + \pi_{h,a}^7 \right)$$

$$- \sum_{a\in\mathcal{A}^{\mathrm{t}}(s,e)} \pi_a^8 \lambda_a - \sum_{t\in\mathcal{T}(l_{\mathrm{r}},s+1,e)} \pi_t^9 - \sum_{t\in\mathcal{T}(l_{\mathrm{r}},s,e)} \pi_t^{10} + \sum_{i\in\mathcal{V}} \pi_i^{11}\underline{x}_i.$$

Note that only one cut is generated at each iteration. One can also use a multi-cut version, by which $|\mathcal{K}|$ Bender's optimality cuts are generated at each iteration (Birge and Louveaux 2011).

E.8. Algorithm 3 Details

Algorithm 3 Heuristic Search ($\mathcal{D}_1$)

```
Initialize $v_{\min} := \infty$, $x_{\text{sum}} := \infty$, $\underline{x}_{\text{sum}} := -\infty$, $\bar{x}_{\text{sum}} := \infty$, $\mathcal{V}' \leftarrow \emptyset$
$\hat{\mathcal{D}}_1 \leftarrow \{(q,v) \mid (\mathbf{x},v) \in \mathcal{D}_1,\ q = \sum_{i\in\mathcal{V}} x_i\}$
$\hat{\mathcal{D}}_1 \leftarrow \{(x_l,v_l) \in \hat{\mathcal{D}}_1 \mid x_1 \leq \cdots \leq x_{|\hat{\mathcal{D}}_1|}\}$, $\bar{x}_{\text{sum}} \leftarrow \hat{\mathcal{D}}_1[|\hat{\mathcal{D}}_1|-1].key$, $\underline{x}_{\text{sum}} \leftarrow \hat{\mathcal{D}}_1[0].key$
$\mathcal{D}_1 \leftarrow \{(\mathbf{x}_l,v_l) \in \mathcal{D}_1 \mid v_1 \leq \cdots \leq v_{|\mathcal{D}_1|}\}$, $v^0 \leftarrow \mathcal{D}_1[0].value$, $\mathbf{x}^0 \leftarrow \mathcal{D}_1[0].key$, $x_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$
$I \leftarrow 0$
if $x_{\text{sum}} \geq \bar{x}_{\text{sum}}$ then
    while $v^0 < v_{\min}$ and $I \leq 20$ do
        Randomly pick $\mathcal{V}' \subseteq \mathcal{V}$ such that $|\mathcal{V}'| = \bar{N}_1$
        $x_i^0 \leftarrow x_i^0 + 1,\ i \in \mathcal{V}'$, $\bar{x}_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v_{\min} \leftarrow v^0$
        Resolve $\mathcal{M}(0,T)$ where $\mathbf{x} \in \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \mathbf{x} = \mathbf{x}^0\}$ to obtain $v^0$
        $\mathcal{D}_1 \leftarrow \mathcal{D}_1 \cup \{(\mathbf{x}^0, v^0)\}$, $I \leftarrow I + 1$
    if $I \leq 20$ then $x_i^0 \leftarrow x_i^0 - 1,\ i \in \mathcal{V}'$, $x_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v^0 \leftarrow v_{\min}$, $v_{\min} \leftarrow \infty$
        while $v^0 < v_{\min}$ and $x_{\text{sum}} + \bar{N}_2 < \bar{x}_{\text{sum}}$ do
            Randomly pick $\mathcal{V}' \subseteq \mathcal{V}$ such that $|\mathcal{V}'| = \bar{N}_2$
            $x_i^0 \leftarrow x_i^0 + 1,\ i \in \mathcal{V}'$, $x_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v_{\min} \leftarrow v^0$
            Resolve $\mathcal{M}(0,T)$ where $\mathbf{x} \in \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \mathbf{x} = \mathbf{x}^0\}$ to obtain $v^0$
            $\mathcal{D}_1 \leftarrow \mathcal{D}_1 \cup \{(\mathbf{x}^0, v^0)\}$
            if $x_{\text{sum}} + \bar{N}_2 < \bar{x}_{\text{sum}}$ then $x_i^0 \leftarrow x_i^0 - 1,\ i \in \mathcal{V}', v^0 \leftarrow v_{\min}$
else
    if $x_{\text{sum}} \leq \underline{x}_{\text{sum}}$ then
        while $v^0 < v_{\min}$ and $I \leq 20$ and $x_i^0 > 0,\ i \in \mathcal{V}$ do
            Randomly pick $\mathcal{V}' \subseteq \mathcal{V}$ such that $|\mathcal{V}'| = \bar{N}_1$
            $x_i^0 \leftarrow x_i^0 - 1,\ i \in \mathcal{V}'$, $\underline{x}_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v_{\min} \leftarrow v^0$
            Resolve $\mathcal{M}(0,T)$ where $\mathbf{x} \in \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \mathbf{x} = \mathbf{x}^0\}$ to obtain $v^0$
            $\mathcal{D}_1 \leftarrow \mathcal{D}_1 \cup \{(\mathbf{x}^0, v^0)\}$, $I \leftarrow I + 1$
        if $I \leq 20$ then $x_i^0 \leftarrow x_i^0 + 1,\ i \in \mathcal{V}'$, $x_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v^0 \leftarrow v_{\min}$, $v_{\min} \leftarrow \infty$
            while $v^0 < v_{\min}$ and $x_{\text{sum}} - \bar{N}_2 > \underline{x}_{\text{sum}}$ and $x_i^0 > 0,\ i \in \mathcal{V}$ do
                Randomly pick $\mathcal{V}' \subseteq \mathcal{V}$ such that $|\mathcal{V}'| = \bar{N}_2$
                $x_i^0 \leftarrow x_i^0 - 1,\ i \in \mathcal{V}'$, $x_{\text{sum}} \leftarrow \sum_{i\in\mathcal{V}} x_i^0$, $v_{\min} \leftarrow v^0$
                Resolve $\mathcal{M}(0,T)$ where $\mathbf{x} \in \mathcal{X} \cap \{\mathbf{x} \in \mathbb{Z}_+^{|\mathcal{V}|} \mid \mathbf{x} = \mathbf{x}^0\}$ to obtain $v^0$
                $\mathcal{D}_1 \leftarrow \mathcal{D}_1 \cup \{(\mathbf{x}^0, v^0)\}$
                if $x_{\text{sum}} - \bar{N}_2 > \underline{x}_{\text{sum}}$ then $x_i^0 \leftarrow x_i^0 + 1,\ i \in \mathcal{V}', v^0 \leftarrow v_{\min}$
Return $\mathbf{x}^0$ and $v^0$
```

Algorithm 3 mainly consists of two steps. First, we set the search step of the total number of allocated vehicles x_{sum} as a positive value $\bar{N}_1 \leq N$. If the solution $\mathbf{x}^0$ in $\mathcal{D}_1$ with the smallest objective value (denoted by $v_{\min}$) has the largest (resp. smallest) total number of allocated vehicles x_{sum}, we then generate a new candidate solution by increasing (resp. decreasing) x_{sum} by $\bar{N}_1$. This is done by randomly choosing x_i, $i \in \mathcal{V}$, for $\bar{N}_1$ times, where for each time the value of the chosen x_i is increased (resp. decreased) by 1. Then, we append this new solution and the corresponding objective value of the two-stage problem $\mathcal{M}(0,T)$ to $\mathcal{D}_1$. If the newly added objective is the smallest in $\mathcal{D}_1$, then we similarly generate another new candidate solution. Otherwise, the search in the first step terminates, and we denote by $\bar{x}_{\text{sum}}$ (resp. $\underline{x}_{\text{sum}}$) the current largest (resp. smallest) total number of allocated vehicles. The intuition of this step is as follows. Suppose we observe the objective values in $\mathcal{D}_1$ monotonically decrease as the corresponding total number of allocated vehicles increases (resp. decreases), then a smaller objective can be found highly possibly by further increasing (resp. decreasing) total number of allocated vehicles (note that we do observe this phenomenon in our numerical experiments, see Section 6). If such observation is not found, we immediately output the solution with the smallest objective value in $\mathcal{D}_1$.

Second, we set the search step of x_{sum} as a positive value $\bar{N}_2 \leq \bar{N}_1/2$ and search within the range between $\bar{x}_{\text{sum}} - |\bar{N}_1|$ and $\bar{x}_{\text{sum}}$ (resp. between $\underline{x}_{\text{sum}}$ and $\underline{x}_{\text{sum}} + |\bar{N}_1|$). That is, we generate a new candidate solution by increasing x_{sum} starting from $\bar{x}_{\text{sum}} - |\bar{N}_1|$ (resp. decreasing x_{sum} starting from $\underline{x}_{\text{sum}} + |\bar{N}_1|$) with a step size $\bar{N}_2$. We append the new solution and the associated objective value of model $\mathcal{M}(0,T)$ to $\mathcal{D}_1$. If the sequence of objective values stops decreasing during the search, then we end up with a locally optimal solution $\mathbf{x}^0$ with the smallest objective in $\mathcal{D}_1$. We can also obtain the corresponding second-stage solution. This second step is to refine the search with a smaller step size hoping that a smaller objective can be found.

In the numerical experiments, we set the perturbation for regional allocation $\sigma_i = 1$ for $i \in \mathcal{V}$, $\underline{\sigma} = \bar{\sigma} = 1$ for the total allocation, and the step sizes $\bar{N}_1 = |\mathcal{V}|$ and $\bar{N}_2 = \lfloor |\mathcal{V}|/3 \rfloor$.

Appendix F: Supplement to Section 6

F.1. Performance of Solution Approaches

We compare the CPLEX approach with our solution approach, and report the results in Table F2.

Table F2 Performance of Solution Approaches

Cases			Weekdays				Weekends			
			The CPLEX Approach		Our Solution Approach		The CPLEX Approach		Our Solution Approach	
c_j	C_r	C_p	Profit ($)	Time (s)	Profit ($)	Time(s)	Profit ($)	Time (s)	Profit ($)	Time (s)
0.5	13	0.1	–	10,800	2,177.6	3,216	-255.8	10,800	877.6	3,563
		0.3	–	10,800	2,177.7	3,356	-767.4	10,800	878.6	3,248
		0.5	–	10,800	2,182.7	2,845	-1,279.1	10,800	879.7	3,319
		0.7	–	10,800	2,190.4	2,121	-1,790.7	10,800	879.2	3,282
	15	0.1	–	10,800	2,169.9	3,037	-255.8	10,800	875.8	3,503
		0.3	–	10,800	2,166.9	2,964	-767.4	10,800	876.6	2,581
		0.5	–	10,800	2,172.6	2,885	-1,279.1	10,800	877.5	3,366
		0.7	–	10,800	2,176.5	2,920	-1,790.7	10,800	877.9	2,322
	17	0.1	–	10,800	2,159.4	3,039	-255.8	10,800	875.2	3,186
		0.3	–	10,800	2,157.8	2,106	-767.4	10,800	875.8	2,977
		0.5	–	10,800	2,165.0	2,229	-1,279.1	10,800	876.6	3,012
		0.7	–	10,800	2,167.2	2,901	-1,790.7	10,800	876.8	2,113
0.8	13	0.1	–	10,800	2,116.4	2,172	-255.8	10,800	848.9	2,822
		0.3	–	10,800	2,118.6	2,816	-767.4	10,800	849.6	2,063
		0.5	–	10,800	2,120.1	1,639	-1,279.1	10,800	849.7	2,820
		0.7	–	10,800	2,123.4	2,464	-1,790.7	10,800	849.7	2,089
	15	0.1	–	10,800	2,107.7	2,065	-255.8	10,800	847.7	2,763
		0.3	–	10,800	2,108.1	2,110	-767.4	10,800	848.0	2,804
		0.5	–	10,800	2,110.2	2,348	-1,279.1	10,800	848.1	2,852
		0.7	–	10,800	2,115.6	2,315	-1,790.7	10,800	848.0	2,485
	17	0.1	–	10,800	2,099.9	2,136	-255.8	10,800	847.2	2,775
		0.3	–	10,800	2,098.9	1,486	-767.4	10,800	847.4	2,854
		0.5	–	10,800	2,100.9	2,467	-1,279.1	10,800	847.9	2,778
		0.7	–	10,800	2,104.2	1,615	-1,790.7	10,800	848.0	2,603
1.0	13	0.1	–	10,800	2,078.3	1,489	-255.8	10,800	831.4	2,425
		0.3	–	10,800	2,077.1	2,051	-767.4	10,800	830.9	2,616
		0.5	–	10,800	2,082.9	2,027	-1,279.1	10,800	830.6	2,598
		0.7	–	10,800	2,088.1	2,196	-1,790.7	10,800	831.2	2,610
	15	0.1	–	10,800	2,063.3	2,156	-255.8	10,800	830.3	2,616
		0.3	–	10,800	2,070.1	2,134	-767.4	10,800	829.5	2,626
		0.5	–	10,800	2,061.6	2,190	-1,279.1	10,800	830.1	2,570
		0.7	–	10,800	2,071.9	1,493	-1,790.7	10,800	829.6	1,804
	17	0.1	–	10,800	2,063.7	2,185	-255.8	10,800	830.1	2,578
		0.3	–	10,800	2,060.2	2,121	-767.4	10,800	828.9	2,572
		0.5	–	10,800	2,053.3	1,470	-1,279.1	10,800	829.6	2,550
		0.7	–	10,800	2,063.8	1,442	-1,790.7	10,800	828.9	2,648

We further perform experiments to assess the optimality gap between the profits obtained by our solution approach and by the original model ($\mathcal{M}$). We follow the setting in Section 6.1 (with $c_j = 0.5$, $C_p = 0.5$, and $C_r = 15$) and consider different instances by varying (i) demand data set, i.e., demands on weekdays and weekends, (ii) the number of scenarios $|\mathcal{K}|$, and (iii) operational horizon $\mathcal{T}$. Tables F3 and F4 show the results for the weekdays and weekends, respectively, where $[a,b]$ denotes $\{a, a+1, \ldots, b\}$ with $b > a$.

For each instance, we compare the performance of three approaches: (i) "CPLEX MIP:" solve the original mixed-integer programming (MIP) model ($\mathcal{M}$) via the CPLEX with its default setting and time limit as 24 hours; (ii) "CPLEX LP:" solve the linear programming (LP) relaxation of the model ($\mathcal{M}$) via the CPLEX with its default setting and time limit as two hours; (iii) "Our Solution Approach:" solve the original MIP model ($\mathcal{M}$) via our proposed solution approach in Section 5.2.

For each approach, we report the profit and computational time with respect to each given instance. For "CPLEX MIP," as no instance can be solved to the optimality within the time limit 24 hours, we report the terminating gap provided by the CPLEX, which is defined as the relative gap between the best profit and the best upper bound of the profit that can be obtained when reaching the time limit. For "Our Solution Approach," we define

$$\text{MIP Gap (\%)} = \frac{\text{(the profit by "CPLEX MIP")} - \text{(the profit by "Our Solution Approach")}}{\text{(the profit by "CPLEX MIP")}} \times 100\%,$$

$$\text{LP Gap (\%)} = \frac{\text{(the profit by "CPLEX LP")} - \text{(the profit by "Our Solution Approach")}}{\text{(the profit by "CPLEX LP")}} \times 100\%$$

with respect to each given instance and report their values.

Table F3 Results with Weekday Demands

No. Scenarios	Horizon	CPLEX MIP			CPLEX LP		Our Solution Approach			
$\vert\mathcal{K}\vert$	$\mathcal{T}$	Profit ($)	Time (s)	Terminating Gap (%)	Profit ($)	Time (s)	Profit ($)	Time (s)	MIP Gap (%)	LP Gap (%)
20	[40, 140]	1,035.0	86,400	4.59	1,095.0	107.8	1,007.1	520.6	2.70	8.03
	[30, 150]	1,136.6	86,400	5.06	1,207.2	136.0	1,107.4	281.7	2.57	8.27
	[140, 240]	1,137.9	86,400	4.33	1,198.6	95.2	1,095.5	385.9	3.73	8.60
	[120, 240]	1,348.3	86,400	4.45	1,421.1	121.3	1,298.7	320.9	3.68	8.61
	[50, 210]	2,075.1	86,400	5.21	2,199.0	221.6	2,008.2	424.2	3.22	8.68
	[40, 220]	2,109.9	86,400	5.66	2,248.3	276.6	2,050.6	520.8	2.81	8.79
40	[40, 140]	1,054.0	86,400	5.35	1,121.9	293.6	1,033.0	1,044.4	1.99	7.92
	[30, 150]	1,157.9	86,400	5.38	1,233.1	555.7	1,123.6	683.4	2.96	8.88
	[140, 240]	1,137.5	86,400	4.64	1,200.5	308.3	1,100.2	843.6	3.27	8.35
	[120, 240]	1,339.2	86,400	4.96	1,416.7	397.5	1,289.3	647.0	3.73	9.00
	[50, 210]	2,094.9	86,400	5.66	2,229.1	632.1	2,028.6	950.6	3.16	8.99
	[40, 220]	2,119.4	86,400	6.76	2,279.3	908.3	2,085.3	1,165.7	1.61	8.51
60	[40, 140]	1,047.9	86,400	5.71	1,118.4	683.6	1,029.2	1,537.2	1.78	7.97
	[30, 150]	1,149.8	86,400	5.74	1,227.4	842.9	1,121.5	1,124.3	2.46	8.63
	[140, 240]	1,090.1	86,400	5.10	1,154.7	472.2	1,056.7	1,355.3	3.06	8.49
	[120, 240]	1,289.4	86,400	5.46	1,369.5	724.9	1,250.1	903.6	3.05	8.72
	[50, 210]	2,039.0	86,400	6.33	2,181.8	1,683.1	1,996.1	1,954.3	2.10	8.51
	[40, 220]	2,074.4	86,400	6.84	2,230.8	2,787.3	2,040.2	1,817.0	1.65	8.54
80	[40, 140]	1,046.3	86,400	5.83	1,117.3	898.5	1,027.9	2,128.6	1.75	8.00
	[30, 150]	1,147.7	86,400	5.89	1,226.4	1,558.4	1,120.1	1,494.8	2.40	8.66
	[140, 240]	1,107.8	86,400	5.09	1,172.3	715.2	1,072.5	1,607.3	3.18	8.51
	[120, 240]	1,308.7	86,400	5.29	1,387.0	1,399.9	1,262.0	1,279.3	3.57	9.01
	[50, 210]	2,016.5	86,400	8.41	2,188.9	3,563.3	2,004.2	2,674.1	0.61	8.44
	[40, 220]	2,076.6	86,400	7.63	2,238.3	3,499.8	2,052.5	2,607.7	1.16	8.30
Average	-	1,464.2	86,400	5.64	1,560.9	953.5	1,427.5	1,178.0	2.59	8.52

Table F3 shows that the "CPLEX MIP" approach cannot solve any instances to the optimality, by which the terminating gap is 5.64% on average. In contrast, our solution approach can solve all the instances within around 20 minutes on average. The profit gap between the "CPLEX MIP" and "Our Solution Approach," i.e., the value of MIP Gap (%), is only 2.59% on average. Note that the profit by the "CPLEX MIP" may not be the true optimal profit (which can be larger), we report the profit by the "CPLEX LP" as the upper bound of the optimal profit. The profit gap between the "CPLEX LP" and "Our Solution Approach," i.e., the value of LP Gap (%), is 8.52% on average. Therefore, the optimality gap between the profit by our solution approach and the optimal profit is between 2.59% and 8.52%.

We observe very similar results in Table F4. Note that the LP relaxation of model ($\mathcal{M}$) may not be successfully solved within the time limit two hours for two large instances with $|\mathcal{K}| = 100$ and $\mathcal{T} = [50,210]$ and $[40,220]$. In contrast, our solution approach can solve the original model ($\mathcal{M}$) of these two instances within around $3,000$ seconds and obtain a better (i.e., larger) profit than the "CPLEX MIP" approach, as suggested by the negative value of MIP Gap (%).

Note that the "CPLEX MIP" cannot solve any instances to the optimality and the profit by the "CPLEX LP" can be too larger than the optimal profit. Thus, we further consider smaller instances that can be solved to the optimality by the "CPLEX MIP," by which an exact value of MIP Gap (%) is obtained to assess the opmality gap (see Table F5). When $\mathcal{T} \in \{[40,140],[140,240]\}$, leading to smaller instances, we use the default setting of the CPLEX, i.e., setting its optimality criterion as 0.01%. When $\mathcal{T} \in \{[50,210],[40,220]\}$, leading to larger instances, we use the default setting of the CPLEX, while setting its optimality criterion as 0.5%. We do not set a time limit and allow the "CPLEX MIP" approach to solve the instances until reaching the optimality criterion. Thus, the profit by the "CPLEX MIP" is considered the optimal profit. Table F5 shows that the value of MIP Gap (%) is 3.76% on average.

In summary, the optimality gap between the profit by our solution approach and the optimal profit is around 4%. Thus, our solution approach significantly reduces the computational time while ensuring very high-quality solutions.

Table F4 Results with Weekend Demands

No. Scenarios $\lvert\mathcal{K}\rvert$	Horizon $\mathcal{T}$	CPLEX MIP			CPLEX LP		Our Solution Approach			
		Profit ($)	Time (s)	Terminating Gap (%)	Profit ($)	Time (s)	Profit ($)	Time (s)	MIP Gap (%)	LP Gap (%)
40	[40, 140]	371.8	86,400	9.38	410.7	708.6	366.4	592.1	1.45	10.79
	[30, 150]	446.5	86,400	11.02	500.2	934.7	434.4	416.4	2.70	13.15
	[140, 240]	576.7	86,400	7.66	626.1	269.0	557.2	655.8	3.39	11.00
	[120, 240]	727.6	86,400	8.51	795.0	574.2	703.1	706.4	3.37	11.56
	[50, 210]	849.8	86,400	11.40	952.0	1,413.3	833.6	1,066.8	1.92	12.44
	[40, 220]	890.0	86,400	10.87	992.6	1,935.3	867.0	1,140.0	2.59	12.66
60	[40, 140]	372.4	86,400	9.79	412.7	1,375.0	367.8	958.8	1.23	10.86
	[30, 150]	450.3	86,400	10.43	501.4	1,688.7	440.2	1,037.3	2.24	12.20
	[140, 240]	572.8	86,400	8.88	628.0	391.0	559.6	1,238.0	2.30	10.89
	[120, 240]	716.5	86,400	10.65	797.3	485.6	709.6	1,310.9	0.96	11.00
	[50, 210]	841.8	86,400	12.87	955.3	4,534.1	838.5	1,671.0	0.40	12.23
	[40, 220]	876.3	86,400	13.00	996.2	3,766.4	872.8	1,884.9	0.40	12.39
80	[40, 140]	346.8	86,400	10.66	387.1	2,361.3	345.1	1,342.2	0.48	10.86
	[30, 150]	418.8	86,400	11.86	472.3	3,252.7	413.6	1,510.1	1.24	12.43
	[140, 240]	542.4	86,400	8.95	594.4	543.9	528.9	1,462.9	2.49	11.02
	[120, 240]	680.9	86,400	10.31	755.3	706.6	673.1	1,820.1	1.14	10.88
	[50, 210]	787.5	86,400	13.87	901.5	6,919.0	789.4	2,272.2	-0.23	12.44
	[40, 220]	820.0	86,400	13.99	939.9	5,338.3	821.1	2,363.9	-0.13	12.64
100	[40, 140]	341.1	86,400	11.93	384.8	2,439.8	342.6	1,836.7	-0.45	10.98
	[30, 150]	415.2	86,400	11.81	468.0	3,509.1	410.6	1,837.9	1.13	12.27
	[140, 240]	524.9	86,400	10.42	583.1	866.0	519.8	2,012.1	0.97	10.84
	[120, 240]	670.2	86,400	10.07	741.8	1,120.6	660.9	2,534.3	1.39	10.91
	[50, 210]	774.0	86,400	14.29	-	7,200.0	779.9	3,011.1	-0.76	-
	[40, 220]	805.0	86,400	14.53	-	7,200.0	810.8	2,912.2	-0.72	-
Average	-	617.5	86,400	11.13	672.5	2,480.5	610.2	1,566.4	1.23	11.66

Table F5 Results with Weekday Demands and $\lvert\mathcal{K}\rvert = 1$

Horizon $\mathcal{T}$	Demand Level	CPLEX MIP			Our Solution Approach		
		Profit ($)	Time (s)	Terminating Gap (%)	Profit ($)	Time (s)	MIP Gap (%)
[40, 140]	low	481.0	27.9	0.01	474.1	93.1	1.43
	medium	953.9	139.7	0.01	929.6	88.2	2.56
	high	1,322.2	33.8	0.01	1,262.3	46.6	4.53
[140, 240]	low	569.2	24.2	0.01	556.6	71.0	2.20
	medium	1,014.4	42.3	0.01	966.5	90.4	4.72
	high	1,407.6	32.1	0.01	1,327.0	52.2	5.72
[50, 210]	low	1,014.1	9,624.9	0.50	990.1	65.8	2.36
	medium	1,903.9	38,122.5	0.50	1,859.4	188.3	2.34
	high	2,611.3	18,808.2	0.50	2,468.6	26.4	5.47
[40, 220]	low	1,030.2	8,928.5	0.50	992.5	73.9	3.66
	medium	1,942.6	96,833.5	0.50	1,850.0	64.8	4.77
	high	2,673.5	47,991.0	0.50	2,529.6	39.2	5.38
Average	-	1,410.3	18,384.0	-	1,350.5	75.0	3.76

F.2. Extending Proposition 1 to The General Case

We observe that Proposition 1 still holds for the general case with multiple regions. To show that Part (a) of Proposition 1 holds, we consider each arc $(n_{it}, n_{j,t+l_{ij}}) \in \mathcal{A}^{\mathrm{t}}$ and $(n_{it}, n_{j,t+l_{\mathrm{r}}}) \in \mathcal{A}^{\mathrm{r}}$ sharing the same start region $i \in \mathcal{V}$, end region $j \in \mathcal{V}$, $j \neq i$, and start time $t \in \mathcal{T}$, and find that $\Lambda^{k^*}_{n_{it},n_{j,t+l_{ij}}} \gamma^{k^*}_{n_{it},n_{j,t+l_{\mathrm{r}}}} = 0$ for any $k \in \mathcal{K}$ (i.e., Part (a) holds by 100%). To show Parts (b) and (c) of Proposition 1 hold for the general case, we need to compare C_{r}/α with the optimal number of vehicles relocated by crowdsourcing (denoted by $\tilde{\Lambda}^k_t$) and that by the 3PL (denoted by $\tilde{\gamma}^k_t$) in each period $t \in \mathcal{T}$ and scenario $k \in \mathcal{K}$. As the two-region case and the general case have different settings, we compute the values of C_{r}/α, $\tilde{\Lambda}^k_t$, and $\tilde{\gamma}^k_t$ in the general case as follows.

First, a fixed variable cost α of crowdsourcing is used in Proposition 1, while the incentive function $g(\cdot)$ is nonlinear and concave in the general case. Thus, for each arc $a \in \mathcal{A}^{\mathrm{t}}$, we obtain a simplified increasing rate $\bar{\phi}_a/(\bar{\Lambda}_a - \epsilon)$ of $g(\cdot)$. We further average this rate over all the arcs in $\mathcal{A}^{\mathrm{t}}$, i.e., $(1/|\mathcal{A}^{\mathrm{t}}|)\sum_{a \in \mathcal{A}^{\mathrm{t}}} \bar{\phi}_a/(\bar{\Lambda}_a - \epsilon)$, to represent the value of α in the general case. *Second*, the 3PL relocation operation is more complicated in the general case. For instance, the relocation cost C_{r} is incurred whenever the 3PL is used in the two-region case, while it is incurred only when the 3PL is requested in a period t in the general case. Once requested in t, the

3PL can be further used in the following $\bar{\bar{z}}$ periods without additional costs. Thus, for any $t \in \mathcal{T}$ and $k \in \mathcal{K}$, we let $\tilde{\Lambda}_t^k = \sum_{s=t}^{t+\min\{\bar{\bar{z}}-1,\, T-1-t\}} \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V},\, j\neq i} \Lambda^{k^*}_{n_{is},n_{j,s+l_{ij}}}$ and $\tilde{\gamma}_t^k = \sum_{s=t}^{t+\min\{\bar{\bar{z}}-1,\, T-1-t\}} \sum_{i\in\mathcal{V}} \sum_{j\in\mathcal{V},\, j\neq i} \gamma^{k^*}_{n_{is},n_{j,s+l_{\mathrm{r}}}}$, where $\Lambda^{k^*}_{n_{is},n_{j,s+l_{ij}}}$ and $\gamma^{k^*}_{n_{is},n_{j,s+l_{\mathrm{r}}}}$ are the optimal solutions to ($\mathcal{M}$).

We then compare the values of C_{r}/α, $\tilde{\Lambda}_t^k$, and $\tilde{\gamma}_t^k$. For Part (c) of Proposition 1, we compare C_{r}/α with $\tilde{\gamma}_t^k$ when $z_t^{k^*} = 1$, $\tilde{\gamma}_t^k > 0$ (i.e., the 3PL is used). We have Part (c) holds by 100% on both weekdays and weekends. For Part (b), we compare C_{r}/α with $\tilde{\Lambda}_t^k$ when both the crowdsourcing and 3PL are available. Note that, during any time interval of l_{f} periods, the 3PL operates for at most $\bar{\bar{z}}$ periods (see (2e)). Thus, the 3PL may not be available for some periods. For example, if $z_t^{k^*} = 1$ and $\tilde{\gamma}_s^k > 0$, $s \in [t,\ t + \bar{\bar{z}} - 1]$, then the 3PL cannot be used since $t + \bar{\bar{z}}$ until $t + l_{\mathrm{f}}$. Thus, in a period t, if we see $z_t^{k^*} = 1$, i.e., the 3PL is requested to serve and hence crowdsouringc is not used by Part (a) of Proposition 1, we then move to period $t + l_{\mathrm{f}}$ to continue the comparisons. The result shows that Part (b) holds by 99.73% on weekdays and 100% on weekends.

F.3. Data Processing via ARIMA Models

In contrast to the previous experiments, we re-generate the training and testing samples separately, with each having 120 samples of scenarios for both the weekday and weekend demands. To that end, the data in 2018 is fitted in seasonal ARIMA models. We use 20 weekdays (resp. 8 weekend days) of trip records in each month to fit a time series model of the bimodal (resp. unimodal) pattern. Thus, we have 12 fitted time series models for weekdays (resp. weekend days) in 12 months. Then, each model generates 20 new samples for the corresponding month by forecasting: 10 for training and 10 for testing, leading to 120 scenarios of demand for training and 120 scenarios of demand for testing. To keep the mean values of the weekday and weekend samples the same in each given month, we follow the following three steps: (i) first, scale the sample values; (ii) then, randomly select an arc in the network $\mathcal{G}$ and add one more demand to the sample with a lower mean value; and (iii) do the above step until the mean values are the same.

Endnotes

1. Source: `https://www.propertyshark.com/mason/ny/New-York-City/Maps?map=nyc2`.
2. Similar user behavior for urban commuters in NYC can be found in NYC DOT (2019).
3. The vehicle allocation cost is estimated by dividing 150 dollars (the normal price of a bicycle at Amazon.com) over 300 operational days per year (we exclude some days due to vehicle maintenance and weather conditions).
4. Given that the annual membership costs 15 dollars per month, the revenue is estimated by assuming that each customer has 30 rides per month with 2.5 periods (15 minutes) per ride, leading to $R = 0.2$.
5. According to American Transportation Research Institute (Murray and Glidewell 2019), the average service cost of motor carriers from 2015 to 2018 is 1.67 dollars per mile. We assume that the 3PL relocation charges the same price and covers 9 miles of a route through nine regions, which leads to about 15 dollars per relocation request.
6. This program has a hierarchical reward scheme. One can earn 1-6 points for one crowdsourcing trip depending on the origin and destination. The value of 1 point also varies. When a rider has accumulated points less than 80 in one month, every 20 points worth a one-week membership (i.e., 15 dollars per month). Thus we round the value of one point as 0.2 dollar. When more points are collected, every 10 points worth 1.2 dollar and we simply take one point as 0.1 dollar. Thus, we assume that on average, a 6-minute trip (i.e., one period) will reward the rider 1-1.5 point (i.e., 0.2-0.3 dollar) and every additional 6-minute will increase the range bounds by 1 more point (i.e., 0.1 dollar). In summary, the reward cost corresponding to trip duration l_{ij} is between $\underline{u}_{ij} = 0.1 + 0.1 l_{ij}$ and $\bar{u}_{ij} = 0.2 + 0.1 l_{ij}$ for any $i, j \in \mathcal{V}$.
7. We also implement the benchmark problems in the recent version CPLEX 12.10, but the solution performance only slightly improved for our instances. Note that the presolver, the dynamic search, the various cuts, and heuristics are turned on by default and decided by CPLEX internally.
8. The relative error is defined as $(|x_{\mathrm{sum},|\mathcal{K}|=k} - x_{\mathrm{sum},|\mathcal{K}|=250}|)/x_{\mathrm{sum},|\mathcal{K}|=250} \times 100\%$, where $x_{\mathrm{sum},|\mathcal{K}|=250}$ and $x_{\mathrm{sum},|\mathcal{K}|=k}$ represent the numbers of allocated vehicles under the 250 scenarios and k scenarios respectively. To compute $x_{\mathrm{sum},|\mathcal{K}|=250}$, we implement our solution algorithm in a single thread by handling each scenario of the second-stage problem ($\mathcal{Q}$) sequentially (rather than in parallel), which requires much less memory.
9. For example, Mobike provides both the 'Lite' version bike (which incurs low cost) and the electric bike (which incurs high cost). The corresponding yearly allocation costs are estimated as 150, 210, and 300, leading to $c_j \in \{0.5, 0.8, 1.0\}$ for all $j \in \mathcal{V}$, respectively. Such costs are representative for most bikes, e-bikes and e-scooters according to their prices listed on Amazon.com; see `https://www.amazon.com/s?k=bike&ref=nb_sb_noss`.
10. The average marginal cost (in US dollar) of trucking per mile ranges from 1.5-1.8 during 2010-2019 (Murray and Glidewell 2019). We assume that the 3PL runs 9 miles per request and thus the total cost per request is 13.5-16.2. We consider a slightly larger range 13-17.
11. The ridership ranking of the subway stations in NYC is referred to MTA (2020).
12. To leave out the numerical errors, we test with different values of C_{p} in a finer scale and a wider range from 0.1 to 1 with step length 0.1. We observe that the number of vehicles relocated by the 3PL remains decreasing while the demand loss shows an increasing trend.