This article was downloaded by: [2a01:4b00:b902:1e00:c5cf:2c57:8f6a:8dfa] On: 17 July 2023, At: 01:37 Publisher: Institute for Operations Research and the Management Sciences (INFORMS) INFORMS is located in Maryland, USA



Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information: http://pubsonline.informs.org

Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance

Philipp Afèche, Zhe Liu, Costis Maglaras

To cite this article:

Philipp Afèche, Zhe Liu, Costis Maglaras (2023) Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance. Manufacturing & Service Operations Management

Published online in Articles in Advance 03 Jul 2023

. https://doi.org/10.1287/msom.2023.1221

Full terms and conditions of use: <u>https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions</u>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2023, INFORMS

Please scroll down for article-it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance

Philipp Afèche,^a Zhe Liu,^{b,*} Costis Maglaras^c

^a Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada; ^b Imperial College Business School, Imperial College London, London SW7 2AZ, United Kingdom; ^c Graduate School of Business, Columbia University, New York, New York 10027 *Corresponding author

Contact: afeche@rotman.utoronto.ca, () https://orcid.org/0000-0002-2010-6983 (PA); zhe.liu@imperial.ac.uk, () https://orcid.org/0000-0001-7335-3024 (ZL); c.maglaras@gsb.columbia.edu (CM)

Received: February 16, 2018 Revised: January 28, 2022; December 11, 2022 Accepted: March 18, 2023 Published Online in Articles in Advance: July 3, 2023

https://doi.org/10.1287/msom.2023.1221

Copyright: © 2023 INFORMS

Abstract. Problem definition: Motivated by ride-hailing platforms such as Uber, Lyft and Didi, we study the problem of matching riders with self-interested drivers over a spatial network. We focus on the performance impact of two operational platform controlsdemand-side admission control and supply-side repositioning control-considering the interplay with two practically important challenges: (i) spatial demand imbalances prevail for extended periods of time; and (ii) self-interested drivers strategically decide whether to join the network, and, if so, whether to reposition when not serving riders. *Methodology*/ results: We develop and analyze the steady-state behavior of a novel game-theoretic fluid model of a two-location, four-route loss network. First, we fully characterize and compare the steady-state system equilibria under three control regimes, from minimal control to centralized control. Second, we provide insights on how and why platform control impacts equilibrium performance, notably with new findings on the role of admission control: the platform may find it optimal to strategically reject demand at the low-demand location even if drivers are in excess supply, to induce repositioning to the high-demand location. We provide necessary and sufficient conditions for this policy. Third, we derive upper bounds on the platform's and drivers' benefits caused by increased platform control; these are more significant under moderate capacity and significant cross-location demand imbalance. Managerial implications: Our results contribute important guidelines on the optimal operations of ride-hailing networks. Our model can also inform the design of driver compensation structures that support more centralized network control.

Supplemental Material: The e-companion and Supplemental Material are available at https://doi.org/10. 1287/msom.2023.1221.

Keywords: ride-hailing • networks • admission control • repositioning • matching • strategic drivers • demand imbalance

1. Introduction

We are motivated by the emergence of ride-hailing platforms such as Uber, Lyft, Didi, and Via that face the problem of matching supply (drivers) with demand (riders) over a spatial network. We study the performance impact of operational platform controls, focusing on the interplay with two practically important challenges: (i) Significant demand imbalances prevail across network locations for extended periods of time (see Figure 1), so that the natural supply of drivers at a location either falls short of or exceeds the demand for rides originating at this location. These mismatches hurt performance, as they lead to lost demand, drivers idling, and/or drivers repositioning (without serving a rider) from a low- to a high-demand location. (ii) Drivers are self-interested and decide strategically whether to join the network, and, if so, when and where to reposition, trading off the related travel time and cost against their matching (queueing) delay at their current location. These decentralized supply decisions may not be optimal for the overall network.

Flow Imbalances: Example Manhattan. We illustrate the magnitude and duration of the demand imbalances noted above with publicly available data for taxi rides in Manhattan.¹ (We do not have public data for ride-hailing platforms, but they likely experience similar imbalances.) Though the data report censored demand (realized trips), we believe the (uncensored) demand imbalances are likely of the same or even higher order of magnitude as the (censored) flow imbalances.

Figure 1 illustrates the *route-level* realized flow imbalances for two origin-destination pairs in Manhattan, New York City, over all weekdays for one month. We observe a pronounced imbalance of almost one order of magnitude (about $10\times$) in the morning rush hour and about half an order of magnitude (about $3\times$) in the



Figure 1. (Color online) Route-Level Flow Imbalances in Manhattan

Notes. (a) Upper West Side-Midtown West, May 2016. (b) Upper East Side-Midtown East, May 2016

evening rush hour in the reverse direction. Our analysis (not shown here) confirms that (i) these route-level flow imbalances persist after aggregation to the *location level*, and (ii) these substantial route- and location-level imbalances are also statistically significant. Furthermore, it is important to note that imbalance periods typically persist for a couple of hours, in contrast to the typical 10–15-minute trip times between these locations. This suggests that network transients may settle down quickly relative to the imbalance duration, which, in part, motivates our focus on the steady-state fluid model as opposed to the transient process itself.

Operational Controls to Manage Demand Imbalances. Motivated by these observations, we study the value of two *operational* platform controls to manage these demand imbalances: demand-side *admission control* and supply-side *capacity repositioning*. Though financial incentives (prices and wages) and spatial information (on demand and price surges) clearly also play an important role in practice, we hold these levers constant to isolate the effects of operational controls. Admission control allows the platform to accept or reject requests based on origin and destination; in practice, platforms do so both directly and indirectly, through ETA quotation. This nonprice control complements pricing, allowing platforms to regulate demand with less drastic price fluctuations. Admission control also affects the car distribution in the network, both directly and indirectly, via drivers' idling delays at lower-demand locations that, in turn, shape their repositioning incentives.

Repositioning control allows the platform, rather than drivers, to decide when and whether they relocate from lower- to higher-demand locations. In practice, such centralized repositioning control is characteristic when drivers operate like employees (e.g., when driving for Via in "Blue Mode" for hourly pay) and will also gain in relevance with the proliferation of autonomous vehicles.

To evaluate these controls, we study the steady-state behavior of a deterministic fluid model of a ride-hailing network in a game-theoretic framework with riders, drivers, and the platform. We provide analytical results for two-location networks (Figure 2) and show through





numerical results for three-location ring and four-location star networks that our main findings generalize to multilocation networks. Riders generate demand for each route (with a fixed travel time). Prices are fixed; we assume for simplicity that the price per unit travel time is route-independent, though this is not necessary for our analysis. Drivers decide, based on their (heterogeneous) opportunity cost and their equilibrium expected profit rate from participation, whether to join the network, and, if so, whether to wait for a rider at their location or to reposition to the other location. Drivers have homogeneous transportation costs and behave symmetrically if they join the network. The platform receives a fixed commission of the fare paid by riders and seeks to maximize its revenue.

We consider three control regimes: (i) *Centralized Control* of both admission and repositioning; (ii) *Minimal Control*, that is, no admission control and decentralized repositioning control; and (iii) *Optimal Admission Control* with centralized admission control and decentralized repositioning control.

Main Results and Contributions. First, we propose a novel game-theoretic model that accounts for key features of ride-hailing platforms: the network structure and demand imbalances, the driver incentives, and the interplay of queueing, transportation times, and driver decisions.

Second, we fully characterize the steady-state system equilibria for the three control regimes outlined above, relying on the analysis of equivalent capacity allocation problems.

Third, we provide insights on how and why platform control impacts equilibrium performance. (i) Decentralized repositioning leads to inefficient capacity allocation as a result of excessive driver idling at low-demand locations. (ii) Admission control can significantly reduce these inefficiencies. (iii) Most notably, we identify a novel role for admission control: as a tool to influence drivers' repositioning decisions. Specifically, the platform may find it optimal to strategically reject demand at the lowdemand location, though there is an excess driver supply, to induce repositioning to the high-demand location. We provide intuitive necessary and sufficient optimal conditions for this policy. This finding highlights that operational levers, and not only pricing, can shape repositioning. Whereas here admission control influences the relocation of idle strategic capacity, the standard roles of admission control in the queueing literature are (1) to balance myopic rewards with opportunity costs, and (2) to control the relocation of *utilized* resources.

Fourth, we derive upper bounds on the platform's and the drivers' benefits caused by increased platform control capabilities. These bounds show that, at practically relevant levels of cross-location demand imbalances, the benefits can be very significant for the platform, of the order of 50%, 100%, or even larger improvements, especially when the network operates with moderate capacity. These bounds also point to tension between platform and driver gains, for example, large platform gains require an increase in driver participation, which limits gains in perdriver profits.

Related Literature. This paper is related to the growing literature on ride-hailing platforms. We first survey theoretical studies and then turn to empirical studies. We group the models considered in theoretical studies into three streams: (i) single-location models with strategic driver supply, (ii) multilocation models with centralized supply control, and (iii) multilocation models with strategic driver supply. Most papers belong to (i) and (ii), whereas this paper belongs to (iii).

Single-Location Models with Strategic Driver Supply. This stream either ignores the spatial dimension or captures it in reduced form. Most studies focus on controlling rider prices and driver wages to match demand with supply; some of these papers largely ignore queueing considerations (e.g., Gurvich et al. 2013, Cachon et al. 2017, and Hu et al. 2022), others consider delay-sensitive customers using queueing models (e.g., Banerjee et al. 2015, Taylor 2018, Bai et al. 2019 and Benjaafar et al. 2021b). Castillo et al. (2016) use a stylized model that captures space in reduced form (pickup times decrease in the number of idle cars) to show that surge pricing can help avoid an inefficient "wild goose chase," whereby long pickup times reduce driver earnings. Garg and Nazerzadeh (2021) study driver surge pricing mechanisms under nonstationary demand. Castro et al. (2021) study priority policies to match drivers with trips that differ in their value.

Multilocation Models with Centralized Supply Control. This stream assumes that platforms fully control the vehicle supply and operation. Most papers model the system as a closed queueing-loss network: nodes correspond to locations, a fixed set of cars circulate among nodes where they queue while waiting for trip matches, and trip requests are lost if not matched upon arrival.

Some papers focus on *demand-side controls*. Waserhole and Jost (2016) and Banerjee et al. (2016) consider static pricing. Balseiro et al. (2021) and Chen et al. (2020) study state-dependent pricing under stationary and nonstationary demand, respectively. Kanoria and Qian (2019) study the joint problem of state-dependent pricing (or admission control) and matching in the absence of prior knowledge of the demand arrival rates. Assuming fixed pricing, Wang et al. (2019) study admission control based on a pickup-time threshold in a two-sided model with open rider-side queue and a closed driver-side queue that captures space in reduced form (similar to Castillo et al. 2016).

In these studies of demand-side control, cars are only matched with local requests, and so only relocate when utilized. In contrast, studies of supply-side controls focus on operational levers to control the flow of empty cars through proactive repositioning and reactive matching. Papers that focus on *repositioning* include those by Iglesias et al. (2016) and Braverman et al. (2019), who study static policies and by Benjaafar et al. (2021) and Hosseini et al. (2021), who consider dynamic policies. Most relevant to our paper is that of Braverman et al. (2019). They prove an asymptotic limit theorem that justifies the use of a stationary deterministic fluid network model (such as the one in this paper) and then characterize the fluidbased optimal empty-car routing policy that maximizes some function of throughput. In contrast to our paper, they fix the capacity, restrict attention to centralized repositioning (as in our regime C in Section 3), and do not consider admission control.

Papers that focus on *matching* include Banerjee et al. (2018), Feng et al. (2020), Özkan and Ward (2020), and Hu and Zhou (2022). Feng et al. (2020) compare the performance of two matching systems, on-demand versus street hailing, for a closed circular queueing network. Banerjee et al. (2018) consider state-dependent control in a closed queueing network; Özkan and Ward (2020) consider state-independent control for an open one-sided queueing model (vehicles exit upon matching); Hu and Zhou (2022) consider dynamic control for a discrete-time, two-sided queueing model (supply and demand units queue before abandoning) and match-dependent rewards reflect spatial distance.

Some papers jointly consider repositioning and matching. An early study by Meyer and Wolfe (1961) compares the performance of various policies for special networks (two nodes, or a continuum of locations with uniformly distributed demand). Ata et al. (2020a) propose and demonstrate the effectiveness of a dynamic policy that hinges on the approximate analysis in the heavy traffic regime.

Some papers study higher-level strategic issues such as capacity sizing (Benjaafar et al. 2021, Besbes et al. 2022) and service region design (e.g., He et al. 2017).

Multilocation Models with Strategic Driver Supply. This stream focuses on pricing policies that account for spatial considerations and strategic drivers' joining and/or location decisions.

Ma et al. (2018) propose an incentive-aligned spatiotemporal pricing mechanism for welfare maximization. Studies of static price and wage policies for revenue maximization include Bimpikis et al. (2019) and Besbes et al. (2021). Bimpikis et al. (2019) consider a discretetime stationary network. They ignore driver queueing effects and assume that ample driver supply is available at a fixed cost, and one-period travel times. They show that platform profits and consumer surplus increase when demands are more balanced across the network, which is consistent with our results that demand imbalances magnify the value of operational controls. Besbes et al. (2021) study short-term location-dependent pricing for a linear city where rational, myopic drivers with exogenous initial locations make one-shot (re)location decisions. Studies of dynamic surge pricing and wage policies under nonstationary demand include Guda and Subramanian (2019) and Afèche et al. (2021).

Unlike these pricing studies, we focus on operational controls. Benjaafar et al. (2021a) adopt our model and extend it by introducing autonomous vehicles (AVs), related operational decisions, and driver wage decisions. They show that if AVs are sufficiently affordable, then the platform would deploy them so as to substantially reduce the need for repositioning by human (strategic) drivers.

Empirical Studies. Some papers study ride-hailing data, others taxi data. Using Uber data, Chen and Sheldon (2016) show that surge pricing induces drivers to work longer and hence increases efficiency; Hall et al. (2017) find that the driver supply is highly elastic to wage and underlying fare changes, so the per-trip earnings boost of a fare hike is negated by higher driver competition (consistent with our results, as noted above). Yan et al. (2020) review operational matching and dynamic pricing techniques and discuss a dynamic waiting mechanism inspired by Uber.

Using NYC taxi data, Buchholz (2022) and Ata et al. (2023) analyze the dynamic spatial equilibrium with strategic taxi drivers, and study how matching and spatial pricing affect performance. Buchholz (2022) shows that matching technology can improve performance significantly even under optimized pricing, which supports the value of studying the impact of operational controls.

Plan for the Paper. In Section 2, we present the model and problem formulations. In Section 3, we study centralized control, and, in Section 4, the regimes with decentralized repositioning. In Section 5, we present theoretical upper bounds on the performance gains of platform control. In Section 6, we generalize our results to multilocation networks. In Section 7, we offer concluding remarks. (The main proofs are in the e-companion; additional technical details are in the Supplemental Material.)

2. Model and Problem Formulations

We consider a deterministic fluid model of a ride-hailing network in steady state. Braverman et al. (2019) rigorously justify such a fluid model for a stochastic closed queueing network with centralized car control in a "large market regime"; that is, the number of cars *N* and the potential demand rates grow linearly with *N*, holding constant travel times. They prove the process-level and steady-state convergence of the scaled queue length process to a fluid limit as $N \rightarrow \infty$. Their arguments could be adapted to our setting; we focus directly on a set of (motivated) steady-state flow equations.

2.1. Model Primitives

Figure 2 shows the network schematic and the model primitives that we describe in this section.

Network. The network has two locations (nodes), indexed by l = 1, 2, and four routes (arcs), indexed by lk for $l, k \in \{1, 2\}$. We denote by t_{lk} the travel time on route lk and by t the travel time vector. We impose no restrictions on travel times; specifically, we allow $t_{12} \neq t_{21}$, to reflect, for example, different uptown/downtown routes. The travel times are constant and, in particular, independent of the number of drivers that serve demand for the platform. This assumes that the number of drivers has no significant effect on road congestion and transportation delays.

Riders. Riders generate demand for trips. The platform charges a fixed price of p per unit of travel time for all routes. The potential demand rate for route-*lk* trips is Λ_{lk} , and Λ denotes the potential demand rate vector. The platform keeps a portion $\gamma \in (0,1)$ of the total fee as commission, and drivers collect the remainder. Rider requests are lost if not matched instantly with an available car. We assume imbalanced crosslocation demands, $\Lambda_{12} \neq \Lambda_{21}$, and the following, without loss of generality.

Assumption 1 (Demand Imbalance). $0 < \Lambda_{12} < \Lambda_{21}$.

Drivers. Drivers supply capacity to the network. Let *N* be the pool of (potential) drivers, each equipped with one car (unit of capacity). Drivers are self-interested and seek to maximize their profit rate per unit time. They decide whether to join the network, and, if so, decide or are directed by the platform whether to reposition (i.e., travel without a rider) from one location to the other.

Participating drivers incur a common driving cost rate of *c* independent of the car occupancy. While serving riders, drivers earn revenue at rate $\overline{\gamma}p$, where $\overline{\gamma} = 1 - \gamma$, and hence profit rate $\overline{\gamma}p - c$. We assume that $\overline{\gamma}p - c > 0$. Drivers' actual profit rate is lower while they wait for riders (zero profit when idling) and/or reposition from one location to the other (incurring the driving cost rate *c*). The following assumption ensures that drivers can earn a positive profit by repositioning.

Assumption 2 (Positive Profit from Repositioning). $ct_{12} < t_{21}(\overline{\gamma}p - c)$ and $ct_{21} < t_{12}(\overline{\gamma}p - c)$.

Each (potential) driver has an idiosyncratic opportunity cost rate, denoted by c_o , that is assumed to be an independent draw from a common continuous distribution *F*.

Assumption 3 (Opportunity Cost Distribution). *The cum*ulative distribution function *F* strictly increases on $[0, \overline{c}_o]$, where $\overline{c}_o \ge p - c$, and satisfies F(0) = 0 and $F(\overline{c}_o) = 1$.

Drivers join the network if and only if their expected profit rate, denoted by π , equals or exceeds their opportunity cost rate. Assumption 3 implies that, given $\pi \in [0, \overline{\gamma}p - c]$, the number of participating drivers $n = NF(\pi) \in [0, N)$. In turn, the per-driver profit rate π emerges in equilibrium and depends on n, the platform's controls,

and the drivers' decisions, as specified in Sections 3 and 4.

Platform. The platform is operated by a monopolist firm that matches drivers with riders with the objective of maximizing its revenue rate. The platform may have two controls: (a) demand-side admission control, and (b) supply-side capacity repositioning, as detailed in Section 2.3.

Information. Riders and drivers rely on the platform for matching; that is, potential riders cannot see the available driver capacity, and drivers cannot see the arrivals of rider requests.

The platform knows the model primitives, including the potential demand rates Λ , the destination of each trip request, the travel times *t*, the driving cost *c*, and the opportunity cost rate distribution *F*. The driver opportunity cost rates are private information, not known by the platform. Therefore, participating drivers are homogeneous to the platform. The platform knows the state of the system, namely, each driver's location, travel direction, and status at each point in time.

Drivers do not observe the system state, but they have (or can infer) the information required to compute their expected profit rates—namely, the travel times *t*; the steady-state delays until they get matched at each location; the destination (routing) probabilities for matches at each location; and the probabilities that they choose or are instructed to reposition from one location to the other. These delays and the routing and repositioning probabilities are endogenous, as detailed below.

2.2. Matching Supply with Demand

Admission Control. Let $\lambda_{lk} \leq \Lambda_{lk}$ denote the *effective* routelk demand rate, that is, the rate of *served* trip requests, and λ the corresponding vector. A trip request is lost if there is no available driver capacity at the time and location of the request, or if the platform exercises admission control to reject the request (e.g., based on the requested destination), even though driver capacity is available.

Matching at Each Location. At each location, drivers that become available (i.e., do not reposition upon arrival) join a single queue, to be matched with accepted ride requests that originate at this location. The platform matches drivers according to a uniform policy, such as first-in-first-out (FIFO) or random order. Therefore, in steady state, drivers queueing at location *l* have the same waiting time, denoted by w_l , and the same matching probability for a route-*lk* request, $\frac{\lambda_{lk}}{\lambda_{l1}+\lambda_{l2}}$. Let q_l denote the steady-state queue length at location *l*. Little's Law implies that $q_l = w_l(\lambda_{l1} + \lambda_{l2})$. Let *w* and *q* denote, respectively, the vector of steady-state waiting times and queue lengths.

Repositioning of Capacity Between Locations. Let v_{12} and v_{21} be the aggregate flow rates of drivers repositioning from location 1 and 2, respectively, and let $v = (v_{12}, v_{21})$.

Up to three flows emanate from location *l*: drivers that are matched with riders leave at rates λ_{l1} and λ_{l2} , and drivers that reposition to location $k \neq l$ leave at rate ν_{lk} (without queueing at location *l*). Therefore, letting $\eta(\lambda, \nu)$ denote the corresponding vector of steady-state repositioning fractions, we have

$$\eta_{1}(\lambda,\nu) = \frac{\nu_{12}}{\lambda_{11} + \lambda_{12} + \nu_{12}} \text{ and} \eta_{2}(\lambda,\nu) = \frac{\nu_{21}}{\lambda_{21} + \lambda_{22} + \nu_{21}}.$$
 (1)

Repositioning decisions are either centralized or decentralized. Under centralized repositioning, the platform controls the repositioning rates ν (e.g., drivers are employees or autonomous vehicles) and the fractions η emerge in response through (1). Under decentralized repositioning, each participating driver chooses his or her repositioning strategy to maximize his or her steadystate profit rate. A driver's repositioning strategy is a vector of probabilities, denoted by $\tilde{\eta}$, that specify for each location the fraction of times that the driver will, upon arrival, directly reposition to the other location. The steady-state profit rate of an individual driver, denoted by $\tilde{\pi}(\tilde{\eta}; \lambda, w)$ and derived explicitly in Section 4.1, is a function of his or her repositioning fractions, $\tilde{\eta}$, the routing probabilities implied by λ , and the delays in the matching queues, w. (The rates λ and delays w, in turn, emerge as equilibrium quantities, as discussed below.) Since participating drivers are homogeneous, we focus on symmetric strategies where drivers choose the same fractions $\tilde{\eta}$ to maximize $\tilde{\pi}(\tilde{\eta}; \lambda, w)$. The flow rates (λ, v) and delays w admit a symmetric driver repositioning equilibrium if, and only if, the resulting unique repositioning fractions $\eta(\lambda, \nu)$ that satisfy (1) agree with every driver's best response to (λ, w) :

$$\eta(\lambda, \nu) \in \underset{\tilde{\eta}}{\operatorname{argmax}} \tilde{\pi}(\tilde{\eta}; \lambda, w).$$
(2)

Steady-State System Flow Constraints. The effective demand rates λ , repositioning flow rates v, waiting times w, and participating driver capacity n must satisfy: (i) the flow balance constraint $\lambda_{12} + v_{12} = \lambda_{21} + v_{21}$; (ii) the capacity constraint $\sum_{l,k=1,2} \lambda_{lk} t_{lk} + (v_{12}t_{12} + v_{21}t_{21}) + \sum_{l=1,2} w_l (\lambda_{l1} + \lambda_{l2}) = n$, where the left-hand side sums the average number of drivers serving riders $(\sum_{l,k=1,2} \lambda_{lk} t_{lk})$, repositioning $(v_{12}t_{12} + v_{21}t_{21})$ and queueing in each location $(\sum_{l=1,2} w_l (\lambda_{l1} + \lambda_{l2}))$.

2.3. Three Control Regimes: Problem Formulations

We study three regimes, *Centralized Control, Minimal Control*, and *Admission Control*, that differ in whether (i) repositioning decisions are centralized or decentralized, and (ii) the platform exercises admission control or not. The problems for all regimes have in common the platform's objective function and drivers' participation decisions that we formalize in this section, as well as the system flow constraints described in Section 2.2. We formulate these problems in terms of the tuple (λ , ν , w, n).

Platform Revenue. Let $\Pi(\lambda) := \gamma p \lambda \cdot t$ denote the platform's steady-state revenue rate, where γp is its commission rate per busy driver and $\lambda \cdot t$ is the number of busy drivers (and riders in service). Hence the realized demand (rider welfare) is proportional to the platform revenue.

Drivers' Participation Constraint. Each driver decides whether to participate by comparing his or her opportunity cost to his or her profit rate from joining the system. We compute the per-driver profit rate in two ways: (i) as the profit rate of an individual driver circulating in the network, $\tilde{\pi}(\tilde{\eta}; \lambda, w)$, as outlined in Section 2.2 and formalized in Section 4.1 (Lemma 2), and (ii) as the average of total driver profits:

$$\pi(\lambda,\nu,n) := \frac{(\overline{\gamma}p - c) \sum_{l,k=1,2} \lambda_{lk} t_{lk} - c(\nu_{12}t_{12} + \nu_{21}t_{21})}{n}$$

As explained in Section 2.2, in each regime, all participating drivers have symmetric repositioning fractions $\eta(\lambda, \nu)$ and hence achieve the same profit rate. The approaches (i) and (ii) therefore yield the same profit rate; that is, $\tilde{\pi}(\eta(\lambda, \nu); \lambda, w) = \pi(\lambda, \nu, n)$ for all (λ, ν, w, n) that satisfy the system flow constraints described in Section 2.2. A *participation equilibrium* therefore requires $n = NF(\pi(\lambda, \nu, n))$.

Centralized Control (C). In this benchmark the platform has "maximum" control and solves the following:

(Problem C)

$$\max_{\lambda,\nu,w,n} \quad \Pi(\lambda)$$

$$\lambda_{12} + \nu_{12} = \lambda_{21} + \nu_{21}, \qquad (3b)$$
$$\sum \lambda_{lk} t_{lk} + \nu_{12} t_{12} + \nu_{21} t_{21}$$

$$+\sum_{l=1,2}^{\infty} w_l(\lambda_{l1} + \lambda_{l2}) = n,$$
 (3c)

(3a)

$$0 \le \lambda \le \Lambda, \, \nu \ge 0, \, w \ge 0, \tag{3d}$$

$$\pi(\lambda, \nu, n) = \frac{(\overline{\gamma}p - c) \sum_{l,k=1,2} \lambda_{lk} t_{lk} - c(\nu_{12}t_{12} + \nu_{21}t_{21})}{n},$$
(3e)

$$n = NF(\pi(\lambda, \nu, n)), \tag{3f}$$

where (3b)–(3c) are the system flow constraints and (3e)–(3f) enforce the participation equilibrium.

Minimal Control (M). In this regime, drivers control repositioning and the platform does not exercise demand admission control. Problem M augments Problem C with the driver repositioning equilibrium constraints (1)–(2), and the following constraints that capture the absence

of admission control. First, the platform matches trip requests to drivers in a pro rata (or FIFO) manner. That is, requests originating at the same location have equal, destination-independent, service probabilities:

$$\frac{\lambda_{l1}}{\Lambda_{l1}} = \frac{\lambda_{l2}}{\Lambda_{l2}}, \quad l = 1, 2.$$
(4)

Second, the platform never turns away requests when there are drivers available to serve them. That is, drivers cannot reposition out of a location that has unmet demand, namely,

$$(\Lambda_{l1} + \Lambda_{l2} - \lambda_{l1} - \lambda_{l2})\nu_{lk} = 0, \quad l = 1, 2, k \neq l,$$
 (5)

and demand requests can only be lost at a location where no drivers are waiting, namely,

$$(\Lambda_{l1} + \Lambda_{l2} - \lambda_{l1} - \lambda_{l2})w_l = 0, \quad l = 1, 2.$$
 (6)

In the Minimal Control regime, the platform therefore solves the following:

(Problem M)

$$\max_{\lambda,\nu,w,n} \{\Pi(\lambda) : (1)-(2), (3b)-(3f), (4)-(6)\}.$$
(7)

Admission Control (A). This regime differs from the centralized benchmark only in that repositioning is decentralized, that is, subject to the driver repositioning equilibrium constraints (1)–(2):

(Problem A)
$$\max_{\lambda,\nu,w,n} \{\Pi(\lambda) : (1)-(2), (3b)-(3f)\}.$$

(8)

2.4. Reformulation to Capacity Allocation Problems

It is intuitive and analytically convenient to reformulate the above problems in terms of the driver capacities allocated to serving riders, repositioning (without riders), and queueing for riders.

Let S_{lk} be the offered load, let s_{lk} be the service capacity on route lk, let S and s be the respective vectors, let $\overline{S} = \sum_{lk} S_{lk}$ be the total load, and let $\overline{s} = \sum_{lk} s_{lk}$ be the total service capacity. By Little's Law,

$$S_{lk} = \Lambda_{lk} t_{lk}$$
 and $s_{lk} := \lambda_{lk} t_{lk}, \quad l,k \in \{1,2\}.$ (9)

Let r_{lk} be the capacity repositioning from location l to k, $r = (r_{12}, r_{21})$, and $\overline{r} = r_{12} + r_{21}$, where

$$r_{lk} = v_{lk} t_{lk}, \quad l \neq k, \tag{10}$$

and let q_l be the capacity queueing at location l. Let $q = (q_1, q_2)$ and $\overline{q} = q_1 + q_2$, where

$$q_l = (\lambda_{l1} + \lambda_{l2})w_l, \quad l = 1, 2.$$
(11)

Using (9)–(11), we transform the problems in (λ, ν, w, n) presented in Section 2.3 into equivalent problems in (s, r, q, n). With some abuse of notation, we write the

platform revenue function as $\Pi(s) = \gamma p\overline{s}$ instead of $\Pi(\lambda)$, the per-driver profit functions as $\tilde{\pi}(\tilde{\eta}; s, q)$ instead of $\tilde{\pi}(\tilde{\eta}; \lambda, w)$ and $\pi(s, r, n)$ instead of $\pi(\lambda, v, n)$, and the repositioning fractions in (1) as $\eta(s, r)$ instead of $\eta(\lambda, v)$.

Using (9)–(11) and the definitions of \overline{s} and \overline{r} , the constraints (3b)–(3f) are equivalent to the following:

$$\frac{s_{12} + r_{12}}{t_{12}} = \frac{s_{21} + r_{21}}{t_{21}},$$
 (12a)

$$\overline{s} + \overline{r} + \overline{q} = n, \tag{12b}$$

$$0 \le s \le S, \ r \ge 0, \ q \ge 0,$$
 (12c)

$$\pi(s,r,n) = \frac{(\overline{\gamma}p - c)\overline{s} - c\overline{r}}{n},$$
 (12d)

$$n = NF(\pi(s, r, n)). \tag{12e}$$

2.5. Two-Step Solution Approach

For regime $X \in \{C, M, A\}$ and capacity n, let $C_X(n)$ denote the set of capacity allocations (s, r, q) that satisfy all the constraints, except the driver participation constraints (12d)-(12e). That is, $C_C(n) = \{(s, r, q) : (12a)-(12c)\}, C_M(n)$ $= \{(s, r, q) : (1)-(2), (4)-(6), (9)-(11), (12a)-(12c)\}$, and C_A $(n) = \{(s, r, q) : (1)-(2), (9)-(11), (12a)-(12c)\}$. Therefore, in regime X, the platform's optimization problem and the optimal revenue rate, denoted by Π_X^* , are given by

$$\Pi_X^* := \max_{s, r, q, n} \{\Pi(s) : (s, r, q) \in \mathcal{C}_X(n), (12d) - (12e)\},\$$

$$X \in \{C, M, A\}.$$
 (13)

We solve the platform's capacity allocation problems (13) in two steps as follows.

Step 1: Solve for the platform-optimal allocation (*s*, *r*, *q*) of a *fixed* driver capacity, *n*:

$$\Pi_X(n) := \max_{s, r, q} \{ \Pi(s) : (s, r, q) \in \mathcal{C}_X(n) \}, \quad \text{for } n \in [0, N].$$
(14)

Let $C_X^*(n) := \operatorname{argmax}_{s,r,q} \{\Pi(s) : (s,r,q) \in C_X(n)\}$. Let $(s_X(n), r_X(n), q_X(n)) \in C_X^*(n)$ be a solution of (14) that maximizes the per-driver profit function in (12d), that is $\pi(s_X(n), r_X(n), n) = \max_{(s,r,q) \in C_X^*(n)} \pi(s,r,n)$. Let $\pi_X(n) := \pi(s_X(n), r_X(n), n)$ be the resulting per-driver profit as a function of n.

Step 2: Solve for the corresponding unique equilibrium participating driver capacity, n_X :

Show that the per-driver profit $\pi_X(n)$ is nonincreasing, with $\lim_{n\downarrow 0} \pi_X(n) = \overline{\gamma}p - c$. This, together with Assumptions 2 and 3, ensures that the platform-optimal solution from Step 1 yields a unique equilibrium participating capacity, $n_X \in (0, N)$, as the unique solution of (12e). The per-driver profit $\pi_X(n)$ is continuous in regimes

C and M but may be discontinuous in regime A, so that

$$n_X = NF(\pi_X(n_X)) \text{ for } X \in \{C, M\}$$

and $NF(\pi_A(n_A^+)) \le n_A \le NF(\pi_A(n_A)).$ (15)

Lemma 1 shows that the two-step solution is optimal if the platform and driver incentives are aligned.

Lemma 1 (Sufficient Condition for Optimality of Two-Step Solution). For regime X, the two-step solution (14)–(15) identifies the optimal solution of (13); that is, $n_X^* = n_X$ and $\Pi_X^* = \Pi_X(n_X)$, if the platform-optimal solution from Step 1 is also driver-optimal at every n; that is, the per-driver profit

$$\pi_X(n) = \max_{(s,r,q) \in \mathcal{C}_X(n)} \pi(s,r,n), \text{ for } n \in [0,N].$$
(16)

Condition (16) holds in regimes C (Corollary 1) and M (Corollary 2), but may not hold in regime A, where we extend the two-step approach by also considering the driver-optimal allocation (Lemma 3).

Remark 1. This two-step approach yields key structural results: (1) Step 1 yields clear results on how the platform-optimal capacity allocation and the resulting per-driver profit depend on the capacity, independent of the opportunity cost distribution F. (2) Lemma 1 links the optimality of the two-step solution to the alignment of platform- and driver-optimality, which also suggests how to extend the approach by balancing these two criteria if they are misaligned. (3) Comparing these results across regimes identifies differences in their capacity allocations, efficiency gains caused by centralized repositioning and admission control, and upper bounds on the gains from these controls.

3. Centralized Control (C)

Problem C, given in (3a)–(3f), is equivalent to the capacity allocation problem (13) for regime X = C. We solve (13) in the two steps specified in Section 2.5. Step 1: For fixed capacity, Proposition 1 characterizes the platformoptimal allocation that solves (14). Step 2: Corollary 1 characterizes the resulting unique equilibrium capacity (15) and establishes that it is optimal by Lemma 1.

Proposition 1 (Regime C: Optimal Allocation of Fixed Driver Capacity). *Define the constants*

$$n_{1}^{C} := S - (\Lambda_{21} - \Lambda_{12})t_{21} \quad and \\ n_{2}^{C} := \overline{S} + (\Lambda_{21} - \Lambda_{12})t_{12}.$$
(17)

In regime C, problem (14) *yields the following optimal allocation* (*s*, *r*, *q*) *of the driver capacity n:*

(1) Scarce capacity $(n \le n_1^C)$. All drivers serve riders: $\overline{s} = n; r = 0; q = 0.$

(2) Moderate capacity $(n_1^C < n \le n_2^C)$. Drivers serve riders or reposition from the low- to the high-demand location:

 $\overline{s} + r_{12} = n$, where $r_{12} = t_{12}/(t_{12} + t_{21})(n - n_1^C)$, $r_{21} = 0$; q = 0.

(3) Ample capacity $(n > n_2^C)$. Drivers serve all riders, reposition from the low- to the high-demand location, or wait in queue: $\overline{s} = \overline{S}$; $r_{12} = n_2^C - \overline{S}$, $r_{21} = 0$; $\overline{q} = n - n_2^C$.

Remark 2. The *equilibrium participating capacity* increases in the driver pool size N. The intervals in n of Proposition 1 hence map to intervals in N with respective capacity allocations in *equilibrium*.

Figure 3(a) in Section 4.4 illustrates Proposition 1. Importantly, centralized control makes efficient use of capacity: drivers idle in queue only if capacity is ample to serve all demand, that is, $n > n_2^C$.

The threshold n_1^C is the maximum offered load that can be served *without* repositioning, that is, all local demand and the balanced cross-location demand; destination-based admission control at the high-demand location allows the platform to serve all local requests while rejecting excess demand to the low-demand location. The threshold n_2^C is the minimum capacity needed to serve the total offered load \overline{S} , including the excess crosslocation demand that requires empty-car repositioning.

Substituting \overline{s} and \overline{r} from Proposition 1 into (12d) yields the following per-driver profit function:

$$\begin{aligned} \pi_{C}(n) &= \frac{(\overline{\gamma}p - c)\overline{s} - c\overline{r}}{n} \\ &= \begin{cases} \overline{\gamma}p - c, & \text{zone 1} (n \le n_{1}^{C}), \\ \frac{1}{n}\overline{\gamma}p \left(n_{1}^{C} + (n - n_{1}^{C})\frac{t_{21}}{t_{12} + t_{21}}\right) - c, & \text{zone 2} (n_{1}^{C} < n \le n_{2}^{C}), \\ \frac{1}{n}(\overline{\gamma}p\overline{S} - cn_{2}^{C}), & \text{zone 3} (n > n_{2}^{C}). \end{cases} \end{aligned}$$

$$(18)$$

This profit rate reflects the drivers' utilization profile: in zone 1, they serve riders all the time; in zone 2, they drive all the time, but serve riders only a fraction of the time; in zone 3, they also queue.

Corollary 1 (Regime C: Driver Participation Equilibrium and Optimal Solution). Under the platform-optimal capacity allocation of Proposition 1, the per-driver profit $\pi_C(n)$ is as follows: (i) continuous and decreasing, so yields a unique driver participation equilibrium, $n_C = NF(\pi_C(n_C))$; (ii) driver-optimal, that is, satisfies (16) in Lemma 1, so the two-step solution is optimal: $n_C^* = n_C$ and $\Pi_C^* = \Pi_C(n_C)$.

4. Regimes with Decentralized Repositioning

We characterize the properties of a driver repositioning equilibrium in Section 4.1, and then the equilibria for two regimes, Minimal Control (M) in Section 4.2; and Admission Control in Section 4.3. We summarize the key differences between the equilibria of regimes C, M, and A in Section 4.4.

4.1. Driver Repositioning Equilibrium

Under decentralized repositioning, the flow rates (λ, ν) and delays w admit a symmetric driver repositioning equilibrium if, and only if, the corresponding unique repositioning fractions $\eta(\lambda, \nu)$, in (1), are every driver's best response to (λ, w) , that is, $\eta(\lambda, \nu) \in \operatorname{argmax}_{\tilde{\eta}} \tilde{\pi}(\tilde{\eta}; \lambda, w)$ (see (2)).

Using (9)–(10) to map (λ, ν) to (s, r), we henceforth express the functions $\eta(\lambda, \nu)$ as $\eta(s, r)$.

Remark 3. Under Assumption 1, it is not optimal to reposition from the high-demand location (2) to the low-demand location (1). We therefore focus on driver repositioning equilibria with $v_{21} = r_{21} = 0$, so that $\eta_2(s, r) = \tilde{\eta}_2 = 0$. Using (9) and (11) to map (λ , w) to (s, q), we henceforth express the profit-rate $\tilde{\pi}(\tilde{\eta}; \lambda, w)$ for simplicity as the univariate function $\tilde{\pi}(\tilde{\eta}_1; s, q)$.

Remark 4. Without loss of optimality, we restrict attention to cases with nonzero served cross-traffic demand; in such cases, $\lambda_{21} > 0$ in light of Remark 3 and flow balance, so that $s_{21} > 0$.

Lemma 2 (Per-Driver Profit Rate). Consider a driver who circulates with repositioning fractions $\tilde{\eta}_1 \in [0,1]$ and $\tilde{\eta}_2 = 0$ through a network with offered loads s and queue lengths q. If $\bar{s}_1 := s_{11} + s_{12} = 0$ and $\tilde{\eta}_1 < 1$, then the driver's expected steady-state profit rate $\tilde{\pi}(\tilde{\eta}_1; s, q) = 0$. Otherwise,

$$\tilde{\pi}(\tilde{\eta}_{1};s,q) = \frac{(\overline{\gamma}p - c)T^{s}(\tilde{\eta}_{1};s) - cT^{r}(\tilde{\eta}_{1})}{T^{s}(\tilde{\eta}_{1};s) + T^{r}(\tilde{\eta}_{1}) + T^{q}(\tilde{\eta}_{1};s,q)},$$
(19)

where $T^s(\tilde{\eta}_1; s)$, $T^r(\tilde{\eta}_1)$, and $T^q(\tilde{\eta}_1; s, q)$ are explicit functions (given in the proof) that denote the expected times that the driver spends in steady state serving riders, repositioning, and queueing, respectively, during a cycle between consecutive arrivals to the same location.

The equilibrium definition (1)–(2) is equivalent to the following definition in terms of (s, r, q).

Definition 1 (Repositioning Equilibrium). A capacity allocation (*s*, *r*, *q*) admits a symmetric driver repositioning equilibrium with $\eta_2(s,r) = 0$ if and only if $r_{21} = 0$ and $\eta(s,r)$ is a driver's best response:

$$\eta_1(s,r) = \frac{r_{12}}{s_{11}\frac{t_{12}}{t_{11}} + s_{12} + r_{12}} \quad \text{and} \\ \eta_2(s,r) = \frac{r_{21}}{s_{21} + s_{22}\frac{t_{21}}{t_{22}} + r_{21}} = 0,$$
(20)

$$\eta_1(s,r) \in \underset{\tilde{\eta}_1}{\operatorname{argmax}} \tilde{\pi}(\tilde{\eta}_1;s,q).$$
(21)

Proposition 2 establishes that condition (21) in Definition 1 implies a mapping from service and repositioning capacities to an explicitly defined set of *driver-incentive compatible* queue lengths.

Proposition 2 (Driver-Incentive Compatible Queue Lengths). *A service and repositioning capacity allocation* (*s*, *r*) *admits a symmetric driver repositioning equilibrium with* $\eta_2(s,r) = 0$, if and only if $r_{12}(s) = \frac{s_{21}t_{12}}{t_{21}} - s_{12}$, $r_{21} = 0$ and the *queue lengths q are driver-incentive compatible, that is,*

$$q \in \mathcal{D}(s) := \begin{cases} \{q : q_1 \leq q_1^*(s) + k(s)q_2\} \\ = \{q : \tilde{\pi}(0; s, q) \geq \tilde{\pi}(1; s, q)\}, \\ if \,\overline{s}_1 > 0 = r_{12}(s), so \, \eta_1(s, r) = 0; \\ \{q : q_1 = q_1^*(s) + k(s)q_2\} \\ = \{q : \tilde{\pi}(0; s, q) = \tilde{\pi}(1; s, q)\}, \\ if \,\overline{s}_1, r_{12}(s) > 0, so \, \eta_1(s, r) \in (0, 1); \\ \{q : q_1 = 0\}, and \ \tilde{\pi}(0; s, q) = 0 < \tilde{\pi}(1; s, q), \\ if \,\overline{s}_1 = 0 < r_{12}(s), so \, \eta_1(s, r) = 1, \end{cases}$$

$$(22)$$

l

where $q_1^*(s) > 0$ and k(s) > 0 are explicit functions that are specified in (A.6).

The corresponding repositioning fraction $\eta_1(s, r)$ *is determined by* (20).

Inducing drivers *not* to reposition from location 1, the first case in (22), requires a sufficiently short location-1 queue, such that the pure strategy of never repositioning weakly dominates that of always repositioning ($\tilde{\pi}(0; s, q) \ge \tilde{\pi}(1; s, q)$). Inducing drivers to reposition from location 1 requires one of two conditions: If any location-1 demand is served ($\bar{s}_1 > 0$, the second case in (22)), then drivers reposition a fraction of the time (identified by (20)) if the queues in the two locations make them indifferent between queueing at and repositioning from location 1. If no location-1 demand is served ($\bar{s}_1 = 0$, the third case in (22)), then drivers prefer to reposition from location 1, so $q_1 = 0$. Proposition 2 foreshadows the key role of admission control in shaping repositioning incentives, discussed in Section 4.3.2.

4.2. Minimal Control (M)

Under minimal control, the platform exercises no admission control and drivers control repositioning. Problem M in (7) is equivalent to the capacity allocation problem (13) for regime X = M. We solve (13) in the two steps specified in Section 2.5. Step 1: for fixed capacity, Proposition 3 characterizes the platform-optimal allocation that solves (14). Step 2: Corollary 2 characterizes the resulting unique equilibrium capacity (15) and establishes that it is optimal by Lemma 1.

We first simplify the set $C_M(n) = \{(s, r, q) : (1)-(2), (4)-(6), (9)-(11), (12a)-(12c)\}$, defined in Section 2.5. Using Proposition 2 to substitute (22) for (1)-(2), and (9)-(11) to

translate (4)–(6) into

$$\frac{S_{l1}}{S_{l1}} = \frac{S_{l2}}{S_{l2}}, \quad l = 1, 2, \tag{23}$$

$$(S_{l1} + S_{l2} - s_{l1} - s_{l2})r_{lk} = 0, \quad l = 1, 2, k \neq l,$$
(24)

$$(S_{l1} + S_{l2} - s_{l1} - s_{l2})q_l = 0, \quad l = 1, 2,$$
(25)

we have

$$C_M(n) = \{(s, r, q) : (12a) - (12c), (22), (23) - (25)\}.$$
 (26)

Proposition 3 (Regime M: Unique Feasible Allocation of Fixed Driver Capacity). *Define*

$$n_1^M := n_1^C - \left(1 - \frac{\Lambda_{12}}{\Lambda_{21}}\right) S_{22},$$

$$n_2^M := n_1^M + q_1^*(S), \quad and \quad n_3^M := n_2^C + q_1^*(S), \quad (27)$$

where n_1^C and n_2^C are defined in (17) and $n_1^M < n_1^C < \overline{S} < n_2^C < n_3^M$.

In regime M, problem (14) has the following unique feasible allocation of the driver capacity, $C_M(n)$:

(1) Scarce capacity $(n \le n_1^M)$. All drivers serve riders: $\overline{s} = n; r = 0; q = 0.$

(2) Moderate capacity, no repositioning but queueing $(n_1^M < n \le n_2^M)$. Drivers serve all riders at the low- and a fraction $\frac{\Lambda_{12}}{\Lambda_{21}}$ of riders at the high-demand location, or queue at the low-demand location: $\overline{s} = n_1^M$, where $s_{1k} = S_{1k}$, $s_{2k} = S_{2k} \frac{\Lambda_{12}}{\Lambda_{21}}$ for k = 1, 2; r = 0; $q_1 = n - n_1^M < q_1^*(S)$, $q_2 = 0$.

(3) Moderate capacity, repositioning and queueing $(n_2^M < n \le n_3^M)$. Drivers serve all riders at the low- and more than a fraction $\frac{\Lambda_{12}}{\Lambda_{21}}$ of riders at the high-demand location, reposition from the low- to the high-demand location, or queue at the low-demand location: $\overline{s} > n_1^M$, where $s_{1k} = S_{1k}$ for k = 1, 2; $r_{12} > 0$, $r_{21} = 0$; $q_1 = q_1^*(S)$, $q_2 = 0$.

(4) Ample capacity $(n > n_3^M)$. Drivers serve all riders, reposition from the low- to the high-demand location, or queue at both locations: $\overline{s} = \overline{S}$; $r_{12} = n_2^C - \overline{S}$, $r_{21} = 0$; q > 0 and $q_1 = q_1^*(S) + k(S)q_2$.

These capacity zones map to driver pool intervals (N) with respective *equilibrium* capacity allocations (cf. Remark 2 for regime C). Figure 3(b) in Section 4.4 illustrates Proposition 3. Compared with Centralized Control (Proposition 1), Minimal Control reduces the driver utilization at moderate capacity (zones 2 and 3); that is, both local and cross-location demand is lost at the highdemand location while drivers idle in queue at the lowdemand location: (1) Because the platform *cannot* use admission control to prioritize local rides at the highdemand location, the maximum load it can serve without repositioning is lower than with admission control (regime C), that is, $n_1^M < n_1^C$. (2) Because repositioning is decentralized, drivers reposition from the low-demand location only if the queue there is sufficiently long, namely, in zone 3, where $q_1^* = n_2^M - n_1^M$. Hence, n_3^M , the minimum capacity required to serve the total offered

load, exceeds the corresponding capacity under centralized control, n_2^C , by exactly q_1^* .

Corollary 2 (Regime M: Driver Participation Equilibrium and Optimal Solution). Under the platform-optimal allocation of Proposition 3, the per-driver profit $\pi_M(n)$ is as follows: (i) continuous and decreasing, so yields a unique driver participation equilibrium, $n_M = NF(\pi_M(n_M))$; (ii) driver-optimal, that is, satisfies (16) in Lemma 1, so the two-step solution is optimal: $n_M^* = n_M$ and $\Pi_M^* = \Pi_M(n_M)$.

4.3. Admission Control (A)

In regime A the platform controls demand admission and drivers control repositioning. This regime may yield the following optimal feature: *strategic demand rejection* at the low-demand location, to encourage drivers to reposition to the high-demand location. In Section 4.3.1 we characterize the equilibrium. In Section 4.3.2 we identify conditions for the optimality of strategic demand rejection.

4.3.1. Equilibrium. Problem A in (8) is equivalent to (13) for regime X = A. We solve (13) by extending the two-step approach specified in Section 2.5 as follows. Step 1: For fixed capacity, Proposition 4 characterizes the platform-optimal allocation that solves (14), where $C_A(n) = \{(s, r, q) : (12a) - (12c), (22)\}$. If strategic demand rejection is optimal at some capacity level, then the platform-optimal policy need *not* be driver-optimal; that is, it may violate condition (16) in Lemma 1. Lemma 3 characterizes the unique equilibrium capacity (15) under the policy of Proposition 4. Lemma 3.1 specifies conditions for this equilibrium capacity to be optimal. Otherwise, Lemma 3.2 specifies how to determine the optimal equilibrium by also applying the two steps to the driver-optimal capacity allocation.

Proposition 4 (Regime A: Optimal Allocation of Fixed Driver Capacity). *Define the constants*

$$n_1^A := n_1^C$$
 and $n_3^A := n_2^C + q_1^*(S)$, (28)

where n_1^C and n_2^C are defined in (17) and $n_1^C < \overline{S} < n_2^C$. In regime A, problem (14) yields the following optimal capacity allocation, where the threshold $n_2^A \in (n_1^A, n_3^A)$ is implicitly defined:

(1) Scarce capacity $(n \le n_1^A)$. All drivers serve riders: $\overline{s} = n; r = 0; q = 0.$

(2) Moderate capacity, no repositioning but queueing $(n_1^A < n \le n_2^A)$. Drivers serve all riders except a fraction $1 - \frac{\Lambda_{12}}{\Lambda_{21}}$ from the high- to the low-demand location, and queue at the low-demand location: $\overline{s} = n_1^A$; r = 0; $q_1 = n - n_1^A < q_1^*(s)$, $q_2 = 0$.

(3) Moderate capacity, repositioning, with or without queueing $(n_2^A < n \le n_3^A)$. Compared with zone 2, drivers serve more riders at the high- but possibly fewer riders at the low-demand location, they reposition from the low- to the high-demand location, and may queue at the low-demand location: $\overline{s} > n_1^A$; $r_{12} > 0$, $r_{21} = 0$; $q_1 = q_1^*(s) \ge 0$, $q_2 = 0$.

(4) Ample capacity $(n > n_3^A)$: Drivers serve all riders, reposition from the low- to the high-demand location, or queue at both locations: $\overline{s} = \overline{S}$; $r_{12} = n_2^C - \overline{S}$, $r_{21} = 0$; $q_1 = q_1^*(S) + k(S)q_2$, $q_2 > 0$.

These capacity zones map to driver pool intervals (N) with respective *equilibrium* capacity allocations (cf. Remark 2 for regime C). Figure 3(c) in Section 4.4 illustrates Proposition 4. Compared with Minimal Control (Proposition 3), Regime A improves driver utilization at moderate capacity (zones 2 and 3) in three ways: (1) With admission control, the platform can serve all local demand at the high-demand location without repositioning, like under Centralized Control, hence $n_1^A = n_1^C > n_1^M$. (2) More local demand admitted at the high-demand location makes this location a more profitable destination for drivers, in turn, reducing the wasteful queueing at the low-demand location. (3) Most importantly, in zone 3 the platform may reject rider requests at the low-demand *location* even if it has available drivers, in order to make it less attractive for drivers to queue there and induce them instead to reposition to, and serve more riders at, the high-demand location. Under this policy, the demand served at the low-demand location decreases in the capacity, that is, in zone 3 compared with zone 2. We call this policy strategic demand rejection, as it serves to regulate the incentives of strategic drivers. We elaborate on the rationale and optimality conditions in Section 4.3.2.

Whereas strategic demand rejection may benefit the platform, it may reduce driver profits, as drivers spend less time queueing and more time repositioning, which is costly. In this case, the platform faces the following trade-off: Increase revenue through strategic demand rejection at the expense of lower driver participation because of reduced per-driver profits; or increase revenue with no (or less) strategic demand rejection, to boost per-driver profits and driver participation. Given this potential trade-off between platform- and driver-optimality, Lemma 3 shows how to extend the two-step solution approach outlined in Section 2.5 to characterize the (optimal) equilibrium capacity n_A^* , and provides sufficient conditions for optimality of strategic demand rejection in equilibrium.

Lemma 3 (Regime A: Driver Participation Equilibrium and Optimal Solution). Under the platform-optimal capacity allocation of Proposition 4, the per-driver profit $\pi_A(n)$ is decreasing and yields a unique equilibrium capacity of participating drivers, n_A , which solves

$$NF(\pi_A(n_A^+)) \le n_A \le NF(\pi_A(n_A)), \text{ where } n_A^+ = \lim_{\epsilon \downarrow 0} (n_A + \epsilon).$$
(29)

1. If the per-driver profit $\pi_A(n_A)$ is driver-optimal, that is,

 π

$$x_A(n_A) = \max_{s,r,q} \{ \pi(s,r,n_A) : (s,r,q) \in \mathcal{C}_A(n_A) \},$$
 (30)

and equals the marginal opportunity cost, i.e., $\pi_A(n_A) = F^{-1}(n_A/N)$, the following holds:

(i) The optimal equilibrium capacity $n_A^* = n_A$, $\pi_A^* = \pi_A(n_A)$, and $\Pi_A^* = \Pi_A(n_A)$.

(ii) Strategic demand rejection is optimal if and only if $n_A \in (n_1^A, n_3^A)$ and $s_{11} < S_{11}$ or $s_{12} < S_{12}$.

2. If (30) fails or $\pi_A(n_A) < F^{-1}(n_A/N)$, then $n_A \in (n_1^A, n_3^A)$, and n_A^* is determined as follows.

Determine the optimal capacity allocation without strategic demand rejection for $n \in [n_1^A, n_3^A]$:

$$\widehat{\Pi}_{A}(n) := \max_{s,r,q} \{ \Pi(s) : (s,r,q) \in \mathcal{C}_{A}(n), s_{11} = S_{11}, s_{12} = S_{12} \}.$$
(31)

Let $\hat{\pi}_A(n)$ be the resulting per-driver profit and $\hat{\pi}_A(n) = \pi_A(n)$ for $n \in [0, n_1^A) \cup (n_3^A, N]$.

(i) Equilibrium capacity without strategic demand rejection: The per-driver profit $\hat{\pi}_A(n)$ is continuous, decreasing, and yields a unique equilibrium capacity, \hat{n}_A , that satisfies

$$\hat{n}_A = NF(\hat{\pi}_A(\hat{n}_A)) \text{ and } \hat{n}_A > n_A.$$
 (32)

(ii) The per-driver profit $\hat{\pi}_A(n)$ is driver-optimal; that is, it satisfies (16) in Lemma 1:

$$\hat{\pi}_A(n) = \max_{(s,r,q)\in\mathcal{C}_A(n)} \pi(s,r,n) \quad \text{for } n \in [0,N].$$
(33)

(iii) The optimal equilibrium participating capacity $n_A^* \in [n_A, \hat{n}_A]$, the per-driver profit π_A^* solves $n_A^* = NF(\pi_A^*)$ with $\pi_A(n_A^*) \le \pi_A^* \le \hat{\pi}_A(\hat{n}_A)$, and $\Pi_A^* \ge \max \{\Pi_A(n_A), \hat{\Pi}_A(\hat{n}_A)\}$.

(iv) To determine the optimal solution $n_A^* \in [n_A, \hat{n}_A]$, solve (13) for fixed $n \in [n_A, \hat{n}_A]$:

$$\overline{\Pi}_{A}(n) := \max_{s, r, q} \{ \Pi(s) : (s, r, q) \in \mathcal{C}_{A}(n), (12d) - (12e) \}$$
(34)

Then, $n_A^* = \operatorname{argmax}\{\overline{\Pi}_A(n) : n \in [n_A, \hat{n}_A]\}, \pi_A^* = F^{-1}(n_A^*/N)$ and $\Pi_A^* = \overline{\Pi}_A(n_A^*).$

(v) If $\Pi_A(n_A) > \hat{\Pi}_A(\hat{n}_A)$, then strategic demand rejection is optimal, the optimal equilibrium capacity $n_A^* < \hat{n}_A$, and the optimal policy harms drivers: $\pi_A(n_A^*) < \hat{\pi}_A(\hat{n}_A)$.

The two-step solution approach yields the unique equilibrium capacity, n_A , in (29). (When the platformoptimal allocation involves strategic demand rejection, the per-driver profit function is discontinuous at some capacity, hence the inequalities in (29).)

Lemma 3.1 establishes two sufficient conditions for the optimality of this equilibrium capacity ($n_A^* = n_A$). The corresponding per-driver profit must be driver-optimal and equal the marginal opportunity cost. In this case, no larger capacity can be an equilibrium, so the two-step solution is optimal (part 1(i)). Furthermore, by part 1(ii), strategic demand rejection is optimal if and only if the equilibrium capacity is moderate (see Proposition 4) and there is unserved demand at the low-demand location. In this case, strategic demand rejection is optimal *without reducing driver profits.*

Lemma 3.2 extends the two-step solution approach to characterize the optimal solution if the sufficient conditions of part 1 are not satisfied. (This can only occur in the intermediate capacity zone, (n_1^A, n_3^A)). In this case, capacity levels larger than n_A can be supported in equilibrium, by increasing the per-driver profit. Part 2 characterizes in (31) the platform-optimal capacity allocation without strategic demand rejection, and shows in part 2(ii) that the resulting per-driver profit $\hat{\pi}_A(n)$ is driveroptimal (see (33)). Part 2(i) characterizes the resulting equilibrium capacity \hat{n}_A in (32) and establishes that it exceeds the equilibrium under the platform-optimal allocation ($\hat{n}_A > n_A$); this highlights the trade-off between driver optimality at larger capacity and platform optimality at smaller capacity. Part 2(iii) shows that the optimal equilibrium capacity n_A^* is "sandwiched" between these equilibrium capacities. Part 2(iv) shows in (34) how to determine the optimal equilibrium capacity n_A^* , as the capacity $n \in [n_A, \hat{n}_A]$ that maximizes the platform revenue, subject to the driver participation constraints (12d)-(12e).

In part 2(v), the sufficient optimality condition for strategic demand rejection, $\Pi_A(n_A) > \hat{\Pi}_A(\hat{n}_A)$, states that the platform revenue under the smaller platform-optimal equilibrium capacity exceeds the platform revenue under the larger driver-optimal equilibrium capacity. In this case, *strategic demand rejection reduces driver participation and profits* $(n_A^* < \hat{n}_A \text{ and } \pi_A(n_A^*) < \hat{\pi}_A(\hat{n}_A))$.

4.3.2. Optimal Strategic Demand Rejection to Induce Driver Repositioning. Strategic demand rejection under moderate capacity (zone 3) means that the platform rejects some or all rider requests at the low-demand location (1), even though there is an *excess* supply of drivers. By sacrificing revenue at the low-demand location, the platform incentivizes drivers to reposition to, and generate more revenue at, the high-demand location. Specifically, rejecting rider requests at the low-demand location creates an artificial demand shortage that drivers offset by choosing to reposition more frequently to the highdemand location, rather than joining the queue at the low-demand location; the result is a shorter queue there (the waiting time may increase or decrease). In terms of Proposition 2, rejecting demand at location 1 alters the driver-incentive compatible capacity allocation by reducing the queue-length threshold $q_1^*(s)$, which frees up driver capacity to reposition and serve riders at the highdemand location. By controlling congestion, the platform uses an operational lever, rather than a financial lever, to incentivize drivers to reposition.

Next, we identify optimality conditions for strategic demand rejection in two steps, (i) at fixed participating capacity, and (ii) at the equilibrium capacity.

Optimality of Strategic Demand Rejection at Fixed Capacity Levels. To simplify notation and highlight the structural imbalances, define

$$\rho_1 := \frac{S_{11}}{S_{11} + S_{12}}, \ \rho_2 := \frac{S_{22}}{S_{21} + S_{22}}, \ \tau := \frac{t_{21}}{t_{12}}, \ \kappa := \frac{c}{\overline{\gamma}p} < 1,$$
(35)

where ρ_1 and ρ_2 are the shares of the local demand offered load at location 1 and 2, respectively, τ is the ratio between cross-location travel times, and κ is the ratio of driving cost to drivers' service revenue ("relative driving cost"). Assumption 2 implies that $\kappa < \tau/(1 + \tau)$.

Proposition 5 identifies a necessary and sufficient condition for the optimality of strategic demand rejection at some fixed capacity levels in the moderate-capacity zone (3) that is defined in Proposition 4.

Proposition 5 (Regime A: Optimality of Strategic Demand Rejection at Fixed Capacity Level). Under optimal admission control and decentralized repositioning, it is optimal at moderate capacity, that is, for some $n \in (n_2^A, n_3^A]$, to strategically reject rider requests at the low-demand location so as to induce repositioning to the high-demand location, if and only if the following condition holds:

$$\frac{\Lambda_{12}}{\Lambda_{21}} \frac{1 - \rho_1 \kappa}{1 - \rho_1} < \frac{\tau - (\tau + 1 - \rho_2)\kappa}{1 - \rho_2} \left(\kappa \frac{1 + \tau}{\tau} \frac{\tau + 1 - \rho_2}{\rho_2} - \tau \right).$$
(36)

Condition (36) is necessary and sufficient for strategic demand rejection to be optimal at *some fixed* capacity, namely, $\Pi_A(n) > \hat{\Pi}_A(n)$ for some *n*. However, optimality of strategic demand rejection at the *equilibrium capacity*, namely, $\Pi_A(n_A^*) > \hat{\Pi}_A(\hat{n}_A)$, requires additional conditions that we present in Proposition 6. First, consider the intuition for (36) to hold, assuming $\tau = 1$ for simplicity:

The share of the local demand offered load at the highdemand location, ρ_2 , is not too high:² Under this condition, drivers have a weak incentive to reposition to the high-demand location, as they are likely to get matched there to a rider going to the low-demand location. Therefore, encouraging drivers to reposition requires rejecting demand at the low-demand location. Without local demand at the high-demand location ($\rho_2 = 0$), condition (36) holds regardless of other factors.

The share of the local demand offered load at the lowdemand location, ρ_1 , is not too high:³ Under this condition, drivers have a strong incentive to queue at the lowdemand location as they are likely to be assigned a rider going to the high-demand location. Therefore, encouraging drivers to reposition to the high-demand location requires rejecting demand at the low-demand location.

The cross-location demand imbalance, $\Lambda_{12}/\Lambda_{21}$, is sufficiently large:⁴ More cross-location demand at the highdemand location increases the value of rejecting demand at the low-demand location in order to induce drivers to reposition to the high-demand location.

The relative driving cost, κ , is sufficiently large: When repositioning is expensive, drivers have no incentive to drive empty; therefore, the platform needs to strengthen their incentive to reposition to the high-demand location over queueing at the low-demand location by rejecting demand there.

Optimality of Strategic Demand Rejection at Equilibrium Capacity. For fixed driver capacity, strategic demand rejection may reduce driver profits, as they pay for repositioning but not for queueing. Therefore, even if (36) holds (so strategic demand rejection is optimal at some fixed capacity levels), the platform may be able to increase revenue without rejecting location-1 demand, by increasing driver profits and participation. Proposition 6 identifies intuitive sufficient conditions for strategic demand rejection to be optimal in equilibrium, that is, for Lemma 3, part 1(ii) or 2(v), to hold.

Proposition 6 (Regime A: Sufficient Optimality Conditions for Strategic Demand Rejection). There exists an interval $(\underline{N}, \overline{N})$ with $0 < \underline{N} < \overline{N} < \infty$ and a threshold function $\overline{\rho}_2(N) : (\underline{N}, \overline{N}) \rightarrow (0, \infty)$ such that strategic demand rejection is optimal in equilibrium if the driver pool size $N \in (\underline{N}, \overline{N})$ and the share of the local demand offered load at the high-demand location $\rho_2 \in [0, \overline{\rho}_2(N))$.

Figure 4 illustrates Proposition 6 with a numerical example. Fixing $\rho_1 = 0.75$, $\Lambda_{21}/\Lambda_{12} = 4$, $\tau = 1$, $\kappa = 0.3$, the left panel shows the region in the (N, ρ_2) -parameter space, that is, the combination of driver pool size and local demand offered load share at the high-demand location, that yields optimal strategic demand rejection in equilibrium. The horizontal line $\rho_2 = 0.53$ indicates

the maximum value of ρ_2 implied by condition (36) in Proposition 5 for optimality of strategic demand rejection at *some capacity level*. The right panel presents this optimality region and threshold line in a way that highlights their connection to the local-demand share at the highdemand location, $\Lambda_{22}/(\Lambda_{21} + \Lambda_{22})$, assuming that local trips originating at the high-demand location last about 20% of cross-location trips (i.e., $t_{22}/t_{21} = 0.2$). This panel shows that strategic demand rejection is optimal even when the local-demand share at the high-demand location is relatively high, up to 60% of total demand.

4.4. Graphical Summary of Capacity Allocation Under Regimes C, M, and A

Figure 3 visualizes for the regimes C, M, and A the optimal capacity allocations specified in Propositions 1, 3 and 4, respectively. To make these graphs comparable, we show these allocations as a function of the same equilibrium capacity (n^* on the horizontal axes), though we note that the equilibrium capacities typically differ across regimes (see Proposition 7). (i) For scarce capacity (zone 1), all drivers are busy serving riders in all regimes. (ii) For ample capacity (zones 3 or 4), all riders are served, and all regimes agree again. But, importantly, (iii) in regime C the platform can serve all the demand with less capacity and no queueing; regimes M and A require the buildup of queues (zone 2) to induce driver repositioning. (iv) Admission control (A) allows the platform to increase driver utilization versus M, by prioritizing demand at the high-demand location based on destination, and also by rejecting demand at the low-demand location to boost repositioning.

For each regime, the model primitives have the following effects: the equilibrium driver participation increases in the driver pool (*N*), the total offered load (\overline{S}) (for fixed route ratios), and the revenue rate ($\overline{\gamma}p$), but decreases in the driving cost (*c*) and the demand imbalance ($\Lambda_{21}/\Lambda_{12}$).







Figure 4. (Color online) Proposition 6: Optimal Strategic Demand Rejection in Equilibrium ($\rho_1 = 0.75, \Lambda_{21}/\Lambda_{12} = 4, \tau = 1, \gamma = 0.25, \kappa = 0.3, F \sim U(0, p - c = 2.55)$)



5. The Impact of Platform Controls on System Performance

In this section, we study how platform controls affect equilibrium performance. In Section 5.1, we rank the key metrics; in Section 5.2, we provide upper bounds on the platform's and the drivers' gains from control.

5.1. Ranking of Platform Revenue, Per-Driver Profit, and Driver Capacity

Proposition 7 shows that more control always benefits the platform but may hurt drivers.

Proposition 7 (Ranking of Equilibrium Profits and Capacity). *Define the driver pool thresholds*

$$\begin{split} N_1^M &:= n_1^M / F(\overline{\gamma}p - c), \quad N_1^A &:= n_1^A / F(\overline{\gamma}p - c), \\ N_3 &:= n_3^A / F(\pi_A(n_3^A)) = n_3^M / F(\pi_M(n_3^M)), \end{split}$$

where n_1^M, n_3^M are defined in (27), n_1^A, n_3^A are defined in (28), and $N_1^M < N_1^A < N_3$.

(1) More platform control increases the equilibrium platform revenue: $\Pi_M^* \leq \Pi_A^* \leq \Pi_C^*$, where

$$\begin{aligned} \Pi_{M}^{*} &< \Pi_{A}^{*} \quad i\!f\!f \quad N \in (N_{1}^{M}, N_{3}) \quad and \quad S_{22} > 0, \\ \Pi_{A}^{*} &< \Pi_{C}^{*} \quad i\!f\!f \quad N \in (N_{1}^{M}, N_{3}). \end{aligned}$$

(2) Centralized control maximizes driver participation and per-driver profit rate: $\max\{n_M^*, n_A^*\} \le n_C^*$ and $\max\{\pi_M^*, \pi_A^*\} \le \pi_C^*$, where the inequalities are strict if and only if $N \in (N_1^A, N_3)$.

(3) With decentralized repositioning, admission control affects driver capacity and profits as follows:

(a) No change $(n_A^* = n_M^*, \pi_A^* = \pi_M^*)$, if the driver pool is scarce $(N \le N_1^M)$ or ample $(N \ge N_3)$.



(b) Increase $(n_A^* > n_M^*, \pi_A^* > \pi_M^*)$, if the driver pool is moderate $(N \in (N_1^M, N_3))$ and strategic demand rejection is suboptimal.

(c) Decrease $(n_A^* < n_M^*, \pi_A^* < \pi_M^*)$, if the driver pool is moderate $(N \in (\underline{N}, \overline{N}') \subset (N_1^A, N_3))$, strategic demand rejection is optimal, and the share of the local demand offered load at the high-demand location is below some threshold $(\rho_2 \in [0, \tilde{\rho}_2(N)))$, where $0 < \tilde{\rho}_2(N) < \infty)$.

Three points emerge from Proposition 7. First, platform control improves performance only if the driver pool is moderate ($N \in (N_1^M, N_3)$). Otherwise, all regimes yield full driver utilization if the pool is scarce ($N \le N_1^M$) or serve all demand if the pool ample ($N \ge N_3$).

Second, for moderate driver pool, more platform control ($M \rightarrow A \rightarrow C$) generally improves the platform revenue (part 1) and the driver participation and per-driver profit (parts 2 and 3a).

Third, drivers may be hurt under decentralized repositioning, in that admission control *reduces* their participation and profits (part 3c) when strategic demand rejection is optimal and stronger conditions hold than those in Proposition 6. Specifically, the conditions on the driver pool $(\overline{N}' < \overline{N})$ and the local demand at the high-demand location $(\tilde{\rho}_2(N) < \overline{\rho}_2(N))$ imply that both the availability and the value of additional capacity are so low that the platform prefers to boost revenue through strategic demand rejection, even at the expense of restricting driver participation.

5.2. Upper Bounds on the Gains in Platform Revenue and Per-Driver Profit

Next, we provide upper bounds on the gains from control for the platform and drivers. **Proposition 8** (Upper Bounds on Platform Revenue Gains). *Fix* $N \ge n_3^M = n_3^A$.

(1) Platform revenue gain because of admission control (regime A over M): If (36) is not satisfied,

$$\max_{F(\cdot)} \frac{\Pi_{A}^{*} - \Pi_{M}^{*}}{\Pi_{M}^{*}} \leq \frac{\overline{S}}{n_{1}^{M}} - 1 = \left(\frac{\Lambda_{21}}{\Lambda_{12}} - 1\right) \frac{1}{1 + \frac{1 - \rho_{21}}{1 - \rho_{1}\tau}}.$$
(37)

(2) Platform revenue gain because of centralized repositioning control (regime C over A):

$$\max_{F(\cdot)} \frac{\prod_{C}^{*} - \prod_{A}^{*}}{\prod_{A}^{*}} \le \frac{\overline{S}}{n_{1}^{A}} - 1 = \left(\frac{\Lambda_{21}}{\Lambda_{12}} - 1\right) \frac{1}{1 + \frac{1}{1 - \rho_{1}\tau} + \frac{\rho_{2}}{1 - \rho_{2}\Lambda_{12}}}$$
(38)

These bounds can be approached arbitrarily closely for specific choices of the opportunity cost distribution $F(\cdot)$ (see Supplemental Material S3). The condition $N \ge n_3^M = n_3^A$ requires a sufficiently large driver pool to serve all riders under decentralized repositioning. The upper bound on the gain from admission control (the right-hand side in (37)) is attained if, under minimal control only the demand that does not require repositioning is served, and admission control increases driver participation enough to serve all demand. The upper bound for repositioning control in (38) has a similar interpretation.

The key insight from (37) and (38) is that the potential revenue gains increase in the cross-location demand imbalance ratio. Table 1 highlights that these gains are very substantial at imbalance ratios such as 2 and 5 that are practically very common (see, e.g., Figure 1 for Manhattan).

Proposition 9 (Upper Bound on Per-Driver Profit Gains). Fix $N \ge n_3^M = n_3^A$, and assume that (36) is not satisfied. The per-driver profit gain from admission control (under regime *A* or *C*) satisfies the following:

$$\max_{F(\cdot)} \frac{\pi_A^* - \pi_M^*}{\pi_M^*} = \max_{F(\cdot)} \frac{\pi_C^* - \pi_M^*}{\pi_M^*} \le \frac{1 - \rho_2}{\tau - (1 - \rho_2 + \tau)\kappa}.$$
 (39)

Whereas the bounds on the platform gains in Proposition 8 can only be attained when more control yields repositioning, attaining the bound on the per-driver profit gain in (39) requires the *absence* of repositioning. This contrast points to the following key tension between the drivers' and the platform's gains from control:⁵ if a small change in the per-driver profit increases the number of participating drivers significantly, then the platform may extract significant gains while drivers are only marginally better off; conversely, if a large change in the per-driver profit only yields a small increase in their number, then drivers extract significant gains while the platform does not.

6. Robustness of Results for Multilocation Networks

In this section, we present numerical results for threelocation ring and four-location star networks. These suggest our analytical results for two-location networks are robust and reveal how they generalize to multilocation networks. We illustrate the key points with selected examples and relegate the mathematical formulations and further numerical results to Supplemental Material S2.

To make our discussion precise, define the *net flow* of location (node) *i* as the difference between its total potential demand inflows and outflows; that is, $NetFlow_i = \sum_k \Lambda_{ki} - \Lambda_{ik}$. We call location *i* an inflow node if $NetFlow_i > 0$, an outflow node if $NetFlow_i < 0$, or a balanced node if $NetFlow_i = 0$.

6.1. Three-Location Ring Networks

Figure 5 shows the four possible types of three-location ring networks; these types differ in the net flow configuration of their nodes. We illustrate the key points with representative results⁶ for network type I. We focus on the optimal capacity allocation for *fixed* capacity *n*, as the equilibrium capacity depends on the driver opportunity cost distribution *F* and pool size *N*. For network I, the offered load $\sum_{lk} S_{lk} = 12$. Balanced demand accounts for nine units, and total excess demand for three units to inflow node 1 (two units from node 2, and one unit from node 3). Figure 6 shows the capacity allocation for Admission Control (A), the most interesting regime (see Supplemental Material S2.4 for Minimal Control), for $n \ge 10$, that is, driver utilization below 100%.

1. Compared with Centralized Control, decentralized repositioning reduces the capacity utilization at moderate

Table 1. Upper Bounds in (37) and (38) on Platform Revenue Gain ($t_{lk} = 1$, $\forall lk$, $\Lambda_{12} = \Lambda_{22} = 1$)

Cross-demand imbalance $\left(\frac{\Lambda_{21}}{\Lambda_{12}}\right)$	1	2	5	10								
From admission control (37)	0%	43%	150%	319%								
From centralized repositioning (38)	0%	25%	100%	225%								
(a) Balanced cross-local demand at low-demand location ($\rho_1 = 0.5$)												
Cross-demand imbalance $\left(\frac{\Lambda_{21}}{\Lambda_{12}}\right)$	1	2	5	10								
From admission control (37)	0%	53%	189%	407%								
From centralized repositioning (38)	0%	30%	120%	270%								
(b) Imbalanced cross-local demand at low-demand location ($\rho_1 = 0.25$)												



Figure 5. (Color online) Four Types of Three-Location Ring Networks

Note. The numbers indicate the potential demand rates of arcs and the net flows of nodes.

capacity, as it requires wasteful queueing to motivate drivers to reposition; furthermore, optimal admission control mitigates these losses versus FIFO admission. Under Centralized Control, the minimum capacity to serve the offered load ($\sum_{lk} S_{lk} = 12$) is $n_2^C = 15$ (80% service utilization, 20% for repositioning) but much larger in Regimes A and M, specifically, $n_3^A = n_3^M = 24$ (50% utilization, 12.5% for repositioning and 37.5% for queueing). Comparing regimes A and M, a capacity of n = 12 serves nine demand units in regime A (Figure 6) but fewer than seven units in regime M (Table 1 in Supplemental Material S2.4), corresponding to service utilizations of 75% versus 56%, respectively.

2. Under decentralized repositioning, drivers only reposition from inflow nodes (here, node 1) to outflow nodes (here, nodes 2 and 3), and the buildup of driver queues obeys the following pattern. As the capacity n increases, driver queues appear at nodes in decreasing order of their net flows:

(i) Queues first form at *inflow nodes* to induce repositioning to outflow nodes: for $n \in [10, 20]$, a

queue forms only at the inflow node 1, where for n > 15 it is long enough to induce repositioning.

(ii) Queues then form at *lower-imbalance outflow nodes*, to incentivize further repositioning to outflow nodes with higher imbalance: For $n \in [21, 24]$, the queue ($q_3 > 0$) at the low-imbalance outflow node 3 reduces this node's attractiveness as a repositioning destination, so drivers at the inflow node 1 are encouraged to reposition to the other (high-imbalance) outflow node 2.

(iii) Queues finally also form at the *higher-imbal*ance outflow node (2) at ample capacity (n > 24).

3. Strategic demand rejection may be optimal at inflow nodes: For $n \in [18, 19]$, some demand is rejected at inflow node 1 to boost repositioning to outflow nodes; this shrinks the node-1 queue.

6.2. Star Networks

These observations also generalize to star networks where the hub is the only inflow node. Figure 7 shows an example: the hub node (1) has a net inflow of 6 and

n	service	s 1 1	s 1 2	s 13	s 21	s 31	r12	r13	<i>q</i> 1	q 2	q 3	η 11	η 12	η13
10	9.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	100%	0%	0%
11	9.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	2.00	0.00	0.00	100%	0%	0%
12	9.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	3.00	0.00	0.00	100%	0%	0%
13	9.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	4.00	0.00	0.00	100%	0%	0%
14	9.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	5.00	0.00	0.00	100%	0%	0%
15	9.20	1.00	1.00	1.00	1.10	1.10	0.10	0.10	5.59	0.00	0.00	94%	3%	3%
16	9.55	1.00	1.00	1.00	1.27	1.28	0.27 🛽	0.28	5.90	0.00	0.00	85%	8% 🛽	8%
17	9.90	1.00	1.00	1.00	1.45	1.45	0.45 🗖	0.45	6.19	0.00	0.00	77%	12% 🛽	12%
18	10.26	1.00	0.97	0.96	1.67	1.67	0.70 🗖	0.71	6.33	0.00	0.00	68%	16% 🔳	16%
19	10.65	1.00	0.81	0.79	2.05	2.00	1.24	1.21	5.85	0.00	0.05	51%	25% 🗖	24%
20	11.00	1.00	1.00	1.00	2.00	2.00	1.00	1.00	7.00	0.00	0.00	60%	20% 🗖	20%
21	11.24	1.00	1.00	1.00	2.24	2.00	1.24	1.00	7.30	0.00	0.23	57%	24% 🔳	19%
22	11.48	1.00	1.00	1.00	2.48	2.00	1.48	1.00	7.59	0.00	0.45	55%	27% 🔳	18%
23	11.74	1.00	1.00	1.00	2.74	2.00	1.74	1.00	7.87	0.00	0.66	52%	30%	17%
24	12.00	1.00	1.00	1.00	3.00	2.00	2.00	1.00	8.14	0.00	0.86	50%	33%	17%
25	12.00	1.00	1.00	1.00	3.00	2.00	2.00	1.00	8.52	0.33	1.14	50%	33%	17%

Figure 6. (Color online) Network I: Optimal Capacity Allocation Under Admission Control (A) Regime

Note. Total demand rate = 12, unit travel times, rider price p = 4, commission rate $\gamma = 25\%$, driving $\cos t c = 1$.

	n	service	s11	s 21	s 31	s 41	r12	r13	r14	<i>q</i> 1	q 2	q 3	q 4	η11	η 12	η13	η14
	12	10.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	100%	0%	0%	0%
	14	10.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	4.00	0.00	0.00	0.00	100%	0%	0%	0%
Queue #4 👝 🖌	16	10.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	6.00	0.00	0.00	0.00	100%	0%	0%	0%
	18	10.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	8.00	0.00	0.00	0.00	100%	0%	0%	0%
$^{-3}(4)$	20	10.05	1.00	1.02	1.01	1.02	0.02	0.01	0.02	9.90	0.00	0.00	0.00	99%	0%	0%	0%
t the second sec	22	10.64	1.00	1.21	1.21	1.21	0.21	0.21	0.21	10.72	0.00	0.00	0.00	86%	5%	5%	5%
	24	11.30	1.00	1.43	1.44	1.43	0.43	0.44	0.43	11.39	0.00	0.00	0.00	75%	8%	8%	8%
ЧТеро	26	12.00	1.00	1.67	1.67	1.67	0.67	0.67	0.67	12.00	0.00	0.00	0.00	67%	11%	11%	11%
1 +6	28	12.72	1.00	1.91	1.91	1.91	0.91	0.91	0.91	12.55	0.00	0.00	0.00	59%	14%	14%	14%
repo 🖊 🧃 👌 Queue #1	30	13.37	1.00	2.00	2.19	2.19	1.00	1.19	1.19	13.12	0.13	0.00	0.00	54%	14%	16%	16%
1 3	32	14.00	0.98	2.00	2.51	2.51	1.00	1.51	1.51	13.64	0.34	0.00	0.00	50%	13%	19%	19%
$1 \sim 2$	34	14.66	0.99	2.00	2.84	2.84	1.00	1.84	1.84	14.15	0.52	0.00	0.00	46%	12%	21%	21%
	36	15.23	0.99	2.00	3.00	3.24	1.00	2.00	2.24	14.68	0.71	0.14	0.00	43%	11%	22%	24%
repo	38	15.72	0.95	2.00	3.00	3.77	1.00	2.00	2.77	15.18	0.92	0.40	0.00	41%	10%	21%	29%
	40	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	15.92	1.15	0.69	0.23	40%	10%	20%	30%
	42	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	16.77	1.46	1.08	0.69	40%	10%	20%	30%
$\underline{\text{Queue } \#2}$ $\underline{\text{Queue } \#3}$	44	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	17.61	1.77	1.46	1.15	40%	10%	20%	30%
	46	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	18.46	2.08	1.85	1.62	40%	10%	20%	30%
	48	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	19.31	2.39	2.23	2.08	40%	10%	20%	30%
	50	16.00	1.00	2.00	3.00	4.00	1.00	2.00	3.00	20.15	2.69	2.62	2.54	40%	10%	20%	30%

Figure 7. (Color online) A Star Network and Its Optimal Capacity Allocation Under Admission Control (A)

the spoke nodes (2, 3 and 4) have net outflows of 1, 2 and 3, respectively. As the capacity *n* increases, we observe similar patterns as in ring network I: queues appear at nodes in decreasing order of their net flow $(1 \rightarrow 2 \rightarrow 3 \rightarrow 4)$, and repositioning between the inflow hub node and the outflow spoke nodes helps serve all demand in the end. We also find strategic demand rejection at the inflow hub node under some intermediate values of *n* (*s*₁₁ < *S*₁₁, as colored red in the table).

7. Discussion and Concluding Remarks

We study the performance impact of operational platform controls for ride-hailing networks with strategic drivers under significant demand imbalances. Our equilibrium analyis of a stationary fluid model yields the following key results: (i) Decentralized repositioning leads to inefficient capacity allocation as a result of excessive driver idling at low-demand locations. (ii) Admission control significantly reduces these inefficiencies. (iii) Most notably, we identify a novel role for admission control: as a tool to influence strategic drivers' repositioning decisions via demand rejection at low-demand locations. The practical implication is that admission control must also consider this effect on the distribution of *empty* cars, not only its immediate effect on busy cars. (iv) We provide upper bounds on the platform's and drivers' benefits caused by increased control. These bounds show that these benefits can be very significant and point to tension between platform and driver gains.

An important direction of future research is to study the interplay of financial and operational controls. The following questions regarding variability and information are also important.

7.1. Steady-State Fluid Model

Ride-hailing services face two types of demand variability: (i) nonstationary average demand rates (e.g., per hour) that reflect significant time-of-day patterns, and (ii) stochastic fluctuations around these time-varying rates. Our model simplifies this setting in two ways, (i) by focusing on a "stationary time slice" during which the demand rates are (approximately) constant, and (ii) by ignoring the stochastic fluctuations. Though we make these simplifications for the sake of analytical tractability, we think the resulting steady-state fluid model provides a reasonable approximation, given the following two key features of demand in operational ride-hailing networks. First, the demand (and supply) rates are large, certainly in major metropolitan areas during rush hour, relative to the effects of stochastic fluctuations; for example, the New York City taxi data shows trip rates on the order of hundreds of trips per hour (equivalently, dozens of trips per 10-minute interval; see Figure 1). This provides informal support for approximating the stochastic discrete model by a deterministic fluid model. Second, though intraday variation in demand rates can be significant, the duration of intraday demand regimes is long (e.g., a couple of hours), compared with typical transportation times (e.g., 10–15 minutes). This suggests that each demand regime (e.g., morning rush vs midday vs. evening rush) is long enough for transients to settle down and the system to reach steady state, or at least that the steady state may be a reasonable approximation.

Though our steady-state fluid model ignores demand variability, it is useful, because it provides a reasonable approximation and yields an analytically tractable formulation that generates important and robust structural results. Specifically, our model allows us to characterize key aspects of strategic drivers' equilibrium behavior and the equilibrium capacity allocation. Whereas this behavior would be intractable under time-varying and stochastic demand, we think in such settings our key insights would continue to hold, notably, (i) decentralized repositioning leads to inefficient capacity allocation caused by excessive driver idling, (ii) admission control can significantly mitigate these inefficiencies, and (iii) how and why admission control may involve strategic demand rejection. Specifically, the key driver of these results is the prevalence of substantial demand imbalances, and empirical data show that such demand imbalances are prevalent in urban traffic. Therefore, we think our key insights would be sufficiently robust under time-varying and/or stochastic demand.

7.2. Variability and Information Design

In our stationary fluid model, the equilibrium system state is constant over time. Therefore, information is irrelevant as a control lever: drivers must simply be informed about (or correctly anticipate) the constant equilibrium values of the key variables that affect their profits, namely, the queueing delays and the demand mix at each location.

An interesting direction for future research is to study the role and value of information design when the equilibrium system state fluctuates because of stochastic and/or time-varying demand.

Under stochastic stationary demand, the platform's information design problem is to decide which queuelength information (if any) to share with drivers. The design where drivers do not observe the idle-car queues is close to our model; namely, in equilibrium, drivers make their repositioning decisions upon arrival to each location, these decisions depend on the steady-state average queue lengths at both locations, and repositioning in equilibrium typically involves mixed strategies and the queue lengths equal some indifference thresholds. The case where drivers observe the real-time queue lengths at both locations gives rise to a dynamic network game with competing long-lived and forward-looking strategic agents. The analysis of this game is challenging for several reasons: (i) drivers can make state-dependent decisions; (ii) the relevant system state depends on all drivers' strategies and is multidimensional (queue lengths at each location, plus number of cars traveling on each route and/or vector of their arrival times at destination); (iii) forecasting their expected queue position upon their next trip completion is difficult for drivers (because it depends on the multidimensional state, and its evolution is subject to demand uncertainty); and (iv) fully forward-looking drivers need to optimize beyond their next trip completion. In sum, it seems imperative to simplify this problem, for example, by restricting drivers' strategy space and/or simplifying their information processing so that they act somewhat myopically, for example, by maximizing their payoff until the next trip completion. Under nonstationary demand, the problem is even more intricate. For example, in addition to the aforementioned challenges, drivers now also need to forecast the effects of changing demand rates on the queue length they can expect at the other location.

Acknowledgments

The authors are grateful to Cyndia Yu, Victoria Lin, and Aruna Prasad for excellent research assistance.

Endnotes

¹New York City TLC Trip Record Data. See https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

² For $\tau = 1$ and $\kappa > 0$, the right-hand side of (36) decreases in ρ_2 from $+\infty$ for $\rho_2 = 0$ to $-\infty$ as $\rho_2 \rightarrow 1$.

³ The left-hand side of (36) increases in ρ_1 from $\Lambda_{12}/\Lambda_{21}$ for $\rho_1 = 0$ to ∞ as $\rho_1 \rightarrow 1$, so that condition (36) holds if both local-demand shares, ρ_1 and ρ_2 , are below some threshold.

⁴ The left-hand side of (36) is positive and decreases to zero as Λ_{21} increases from Λ_{12} to ∞ ($\Lambda_{12}/\Lambda_{21} < 1$ by Assumption 1). Therefore, (36) holds for sufficiently large Λ_{21} , provided the right-hand side is positive, that is, ρ_2 is below some threshold.

⁵ Supplemental Material S3 illustrates this tension and how it depends on the opportunity cost distribution $F(\cdot)$.

⁶ Types II and III yield similar results; see Supplemental Material S2. We omit type IV, as it is balanced.

References

- Afèche P, Liu Z, Maglaras C (2021) Surge pricing and dynamic matching for hotspot demand shock in ride-hailing networks. Working paper.
- Ata B, Barjesteh N, Kumar S (2020) Dynamic dispatch and centralized relocation of cars in ride-hailing platforms. Preprint, submitted September 14, https://dx.doi.org/10.2139/ssrn.3675888.
- Ata B, Barjesteh N, Kumar S (2023) Enhancing equitable access to taxis in NYC through search friction reduction and spatial pricing. Preprint, submitted January 18, https://dx.doi.org/10.2139/ssrn.4324581.
- Bai J, So KC, Tang CS, Chen XM, Wang H (2019) Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing Service Oper. Management* 21(3): 556–570.
- Balseiro SR, Brown DB, Chen C (2021) Dynamic pricing of relocating resources in large networks. *Management Sci.* 67(7): 4075–4094.
- Banerjee S, Freund D, Lykouris T (2016) Pricing and optimization in shared vehicle systems: An approximation framework. Preprint, submitted August 24, https://doi.org/10.48550/arXiv. 1608.06819.
- Banerjee S, Kanoria Y, Qian P (2018) Dynamic assignment control of a closed queueing network under complete resource pooling. Preprint, submitted March 13, https://doi.org/10.48550/arXiv. 1803.04959.
- Banerjee S, Riquelme C, Johari R (2015) Pricing in ride-share platforms: A queueing-theoretic approach. Preprint, submitted February 24, https://dx.doi.org/10.2139/ssrn.2568258.
- Benjaafar S, Wang Z, Xiaotang Y (2021a) Autonomous vehicles for ride-hailing. Preprint, submitted September 10, https://dx.doi. org/10.2139/ssrn.391941.
- Benjaafar S, Ding JY, Kong G, Taylor T (2021b) Labor welfare in on-demand service platforms. *Manufacturing Service Oper. Man*agement 24(1):110–124.
- Benjaafar S, Wu S, Liu H, Gunnarsson EB (2021) Dimensioning on-demand vehicle sharing systems. *Management Sci.* 68(2): 1218–1232.
- Besbes O, Castro F, Lobel I (2021) Surge pricing and its spatial supply response. *Management Sci.* 67(3):1350–1367.
- Besbes O, Castro F, Lobel I (2022) Spatial capacity planning. Oper. Res. 70(2):1271–1291.
- Bimpikis K, Candogan O, Saban D (2019) Spatial pricing in ridesharing networks. Oper. Res. 67(3):744–769.
- Braverman A, Dai JG, Liu X, Ying L (2019) Empty-car routing in ridesharing systems. Oper. Res. 67(5):1437–1452.
- Buchholz N (2022) Spatial equilibrium, search frictions and efficient regulation in the taxi industry. *Rev. Econom. Stud.* 89(2):556–591.

18

- Cachon GP, Daniels KM, Lobel R (2017) The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing Service Oper. Management* 19(3):368–384.
- Castillo JC, Knoepfle DT, Weyl EG (2016) Matching in ride hailing: Wild goose chases and how to solve them. Preprint, submitted December 28, https://dx.doi.org/10.2139/ssrn.2890666.
- Castro F, Ma H, Nazerzadeh H, Yan C (2021) Randomized FIFO mechanisms. Preprint, submitted November 21, https://doi. org/10.48550/arXiv.2111.10706.
- Chen MK, Sheldon M (2016) Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform. *EC '16: Proc. 2016 ACM Conf. Economics and Computation*, 455. https:// doi.org/10.1145/2940716.2940798.
- Chen Q, Lei Y, Jasin S (2020) Real-time spatial-intertemporal dynamic pricing for balancing supply and demand in a ride-hailing network. Preprint, submitted June 19, https://dx.doi.org/10.2139/ ssrn.3610517.
- Feng G, Kong G, Wang Z (2020) We are on the way: Analysis of on-demand ride-hailing systems. *Manufacturing Service Oper. Management* 23(5):1005–1331.
- Garg N, Nazerzadeh H (2021) Driver surge pricing. *Management Sci.* 68(5):3219–3235.
- Guda H, Subramanian U (2019) Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. *Management Sci.* 65(5):1995–2014.
- Gurvich I, Lariviere M, Moreno A (2013) Operations in the on-demand economy: Staffing services with self-scheduling capacity. Preprint, submitted October 6, https://dx.doi.org/10.2139/ssrn.2336514.
- Hall JV, Horton JJ, Knoepfle DT (2017) Labor market equilibration: Evidence from Uber. Working paper, http://john-joseph-horton. com/papers/uber_price.pdf.
- He L, Mak HY, Rong Y, Shen ZJM (2017) Service region design for urban electric vehicle sharing systems. *Manufacturing Service* Oper. Management 19(2):309–327.

- Hosseini M, Milner J, Romero G (2021) Dynamic relocations in carsharing networks. Preprint, submitted February 18, https://dx. doi.org/10.2139/ssrn.3774324.
- Hu B, Hu M, Zhu H (2022) Surge pricing and two-sided temporal responses in ride hailing. *Manufacturing Service Oper. Management* 24(1):91–109.
- Hu M, Zhou Y (2022) Dynamic type matching. Manufacturing Service Oper. Management 24(1):125–142.
- Iglesias R, Rossi F, Zhang R, Pavone M (2016) A BCMP network approach to modeling and controlling autonomous mobilityon-demand systems. Preprint, submitted 15 July, https://doi. org/10.48550/arXiv.1607.04357.
- Kanoria Y, Qian P (2019) Blind dynamic resource allocation in closed networks via mirror backpressure. Preprint, submitted March 7, https://doi.org/10.48550/arXiv.1903.02764.
- Ma H, Fang F, Parkes DC (2018) Spatio-temporal pricing for ridesharing platforms. Preprint, submitted January 11, https://doi. org/10.48550/arXiv.1801.04015.
- Meyer RF, Wolfe HB (1961) The organization and operation of a taxi fleet. Naval Res. Logist. 8(2):137–150.
- Özkan E, Ward AR (2020) Dynamic matching for real-time ride sharing. Stochastic Systems 10(1):29–70.
- Taylor TA (2018) On-demand service platforms. Manufacturing Service Oper. Management 20(4):704–720.
- Wang G, Zhang H, Zhang J (2019) On-demand ride-matching in a spatial model with abandonment and cancellation. Preprint, submitted July 8, https://dx.doi.org/10.2139/ssrn.3414716.
- Waserhole A, Jost V (2016) Pricing in vehicle sharing systems: Optimization in queuing networks with product forms. EURO J. Transportation Logist. 5:293–320.
- Yan C, Zhu H, Korolko N, Woodard D (2020) Dynamic pricing and matching in ride-hailing platforms. *Naval Res. Logist.* 67(8):705–724.