Real-Time Spatial-Intertemporal Pricing and Relocation in a Ride-Hailing Network: Near Optimal Heuristics and The Value of Dynamic Pricing

Qi (George) Chen
London Business School, gchen@london.edu

Yanzhe (Murray) Lei
Smith School of Business, Queen's University, yl64@queensu.ca

Stefanus Jasin
Ross School of Business, University of Michigan, sjasin@umich.edu

Motivated by the growth of ride-hailing services in urban areas, we study a (tactical) real-time spatial-intertemporal dynamic pricing problem where a firm uses a pool of homogeneous servers (e.g., a fleet of taxis) to serve price-sensitive customers (i.e., a rider requesting a trip from an origin to a destination) within a finite horizon (e.g., a day). We consider a revenue maximization problem in a model that captures the stochastic and non-stationary nature of demands, and the non-negligible travel time from one location to another location. We first show that the relative revenue loss of *any* static pricing control is at least in the order of $n^{-1/2}$ in a large system regime where the demand arrival rate and the number of servers scale linearly with n, which highlights the limitation of static pricing control. We also propose a static pricing control with a matching performance (up to a multiplicative logarithmic term). Next, we develop a novel state-dependent dynamic pricing control with a reduced relative revenue loss of order $n^{-2/3}$. The key idea is to dynamically adjust the prices in a way that reduces the impact of past "errors" on the balance of future distributions of servers and customers across the network. Our extensive numerical studies using both synthetic and real data set from Manhattan Yellow Taxi confirm our theoretical findings and highlight the benefit of dynamic pricing over static pricing, especially when dealing with non-stationary demands. Interestingly, we also observe that the revenue improvement under our proposed control primarily comes from an increase in the number of customers served instead of from an increase in the average prices compared to the static pricing control. This suggests that dynamic pricing can be potentially used to simultaneously increase both revenue and the number of customers served (i.e., service level). Finally, as an extension, we discuss how to generalize the proposed control to a setting where the firm can also actively relocate some of the available servers to different locations in the network in addition to implementing dynamic pricing.

1. Introduction

Urban mobility is an important topic faced by many cities around the world. One critical part of managing urban mobility is to have an effective ride-hailing system in place. Managing such a system is challenging because urban traffic faces a lot of variability on the demand side. While some of these are *predictable* variabilities due to the non-stationary pattern of demands both over time and across different locations, others are *unpredictable* variabilities due to the randomness of demands (i.e., the difference between the actual realized demands and their expected values).

One typical problem in managing a ride-hailing system is how to dynamically balance supply (drivers or servers) and demand (riders or customers) over time and across different regions. This is usually done using many different levers such as pricing, driver relocation from some regions to other regions, etc. Indeed, the use of dynamic pricing to balance supply and demand has been gaining popularity not only in the context of non-traditional companies such as Uber and Lyft but also in the context of more traditional taxi companies. For example, in 2017, the Land Transport Authority and the Public Transport Council in Singapore approved the proposals put forth by the taxi companies to implement dynamic pricing for taxi fares (Lim 2017). Since then major taxi companies have implemented dynamic pricing systems by either developing their own software infrastructure or partnering with other companies such as Grab and Uber.

In this paper, we focus on both pricing and relocation decisions. Specifically, we consider the setting of a ride-hailing firm with a fixed number of drivers (such as in the traditional taxi companies), non-stationary demands over time and across different regions, non-negligible deterministic travel time from one region to another region, and ask *how can the firm maximize its expected revenues through dynamic pricing and dynamic relocation?* This is technically a hard problem. For example, the pricing decisions at the current period not only affect the immediate number of rides for any pair of origin and destination but also the available number of drivers at the destination point at a future time after the drivers drop off the riders and become available for new service. In other words, the geographic distribution of available supply is *correlated* over time and is affected by the pricing decisions in a non-trivial way, which suggests that an effective dynamic pricing control needs to take such interdependency over time and space into consideration. In this paper, we focus on developing and analyzing provably near-optimal heuristic controls. Due to the complexity of the problem, we first discuss the case where the firm only uses dynamic pricing to balance supply and demand (Sections 3 to 6) and then we discuss the case where the firm uses both dynamic pricing and dynamic relocation (Section 7).

Our results and contributions in this paper can be summarized as follows:

1. We first analyze the performance of a family of *static* pricing controls. (We use the terminology *static control* to refer to the class of open-loop controls, also known as state-independent controls, and the terminology *dynamic control* to refer to the class of closed-loop controls, also known as adaptive controls or state-dependent controls. Note that a static control is allowed to have *non-stationary* control parameters over time, which could be beneficial when the demand pattern is also non-stationary.) We show that the relative revenue loss of *any* static pricing control is at least in the order of $n^{-1/2}$ in a large system regime where the demand arrival rate and the number of servers scale linearly with n. We also propose a static pricing

control, simply called the *Static Price Control* (SPC), with a matching performance (up to a multiplicative logarithmic term).

2. We then develop a novel dynamic pricing control, called the *Arc Balancing Control* (ABC), which guarantees a relative revenue loss of order $n^{-2/3}$ (up to multiplicative logarithmic terms). Central to the design of ABC is a novel batching idea which adaptively adjust the realized variability in demand arrival. Clearly, ABC improves the asymptotic performance of SPC. Moreover, to the best of our knowledge, this is the first heuristic in the ride-hailing literature which establishes a sub-$n^{-1/2}$ revenue loss in a setting with non-stationary demand arrival and positive travel time. Unlike many papers in the ride-hailing literature (see Section 2) that study the *steady-state* control problem, in this paper, we study the *transient* control problem. This is motivated by the fact that demands in ride-hailing system tend to be highly non-stationary. Moreover, as reported by Braverman et al. (2019) in their numerical studies in a closely related setting, depending on the parameters and initial conditions of a ride-hailing system, the convergence of the system to equilibrium could take a relatively long time compared to the length of the whole horizon (about 10 hours when the length of the horizon is one day), which highlights the importance of studying a transient control problem.

3. We conduct extensive numerical simulations using both synthetic data and real data set from Manhattan Yellow Taxi. The numerical results confirm our theoretical findings and highlight the benefit of dynamic pricing over static pricing, with revenue improvements of about 2% - 6%. Moreover, we also observe that the revenue improvement of ABC over SPC comes more from the increase in the number of customers served than from the increase in the average prices (we notice an increase in the number of customers served from about 1% - 4%). This observation provides an interesting insight that dynamic pricing can be used to not only increase revenue but also the number of customers served (i.e., service level), which is also one of the most important goals of an urban transportation system.

4. Motivated by common practices in ride-hailing platforms, we consider an extension where the firm can jointly control the prices of rides and the relocation of servers that are not currently in use. We propose a generalization of ABC, termed R-ABC, that also adaptively adjusts the relocation quantity in a batched manner. We show that the relative revenue loss of R-ABC is also in the order of $n^{-2/3}$, which confirms that the proposed batch adjustment scheme can be applied to broader types of decisions.

The remainder of the paper is organized as follows: We review the most relevant works in the next section. Section 3 formalizes the model and introduces the performance metric. Section 4 investigates the best achievable performance within the class of static pricing controls and proposes SPC, which is near-optimal in the class of static pricing controls. In Section 5, we discuss our main dynamic pricing control, ABC, whose performance strictly improves that of static pricing controls. Section 6 complements the analytical results with numerical studies. We consider a model extension in which the firm can also dynamically relocate servers in addition to implementing dynamic pricing in Section 7. Finally, we conclude the paper in Section 8. Unless otherwise noted, the omitted proofs and further details to the numerical studies can be found in the Appendix (electronic companion) of the paper.

2. Related literature

Our paper is motivated by the ride-hailing systems and is thus closely connected to a growing body of literature on online platforms which provide different kinds of services (e.g., Uber for urban mobility, Upwork for professional freelancer services, Deliveroo for take-away food delivery) to customers using a pool of independent service agents. Some of the early papers that study these platforms focus on the impact of self-scheduling service agents on platform's operational decisions such as pricing (e.g, prices for customers and wages to service agents) and the resulting welfare implications via either analytical models (Cachon et al. 2017, Castillo et al. 2017, Taylor 2018, Chen and Hu 2019, Fang et al. 2019, Bai et al. 2019, Yan et al. 2019, Garg and Nazerzadeh 2021, Guda and Subramanian 2019) or empirical approach (Ata et al. 2019). However, since many papers in this stream of literature focus on strategic market design questions, they tend to abstract away from some features which are unique to ride-hailing systems such as the fact that the number of available drivers across different regions at any given time of the day are not always well-distributed compared to demands. To investigate the impact of this type of inefficiency, a stream of literature has focused on understanding the financial consequence of the spatial imbalance of supply and demand by explicitly modeling the *steady state* equilibrium dynamics of rider and driver flows on a network. Some papers in this literature focus on the class of static pricing controls where pricing decisions are determined a priori (Banerjee et al. 2021, Ozkan 2018, Besbes et al. 2021b, Bimpikis et al. 2019), whereas others focus on developing dynamic pricing controls where prices can be adjusted based on real-time imbalances of demand and supply across the network (Balseiro et al. 2021, Kanoria and Qian 2019, Varma et al. 2022).

Compared to the above stream of papers on ride-hailing systems, our model and analysis are very different. First and foremost, whereas most of the prior papers assume that the system is already in steady state and focus on the steady-state control problem, our model does not require that assumption and we analyze the transient control problem directly. From the practical perspective, whether the assumption that the system is already in steady state is appropriate depends on many factors such as the level of non-stationarity of demand over time (e.g., how fast the demand process changes throughout the course of a day) and how fast the system reaches steady state under any given demand environment (e.g., this potentially depends on many factors such as how much traffic volume there is, the type of controls used, etc.). If, for example, demand pattern varies quickly on an hourly basis and the total hourly volume of rides is not sufficiently large, then one would suspect that the system is more likely spending most of the time in transient state and, as a result, steady state analysis may not provide an accurate depiction of the dynamics of the system. In fact, as reported by Braverman et al. (2019) in their numerical studies in a closely related setting, depending on the parameters and initial conditions of a ride-hailing system, the convergence of the system to equilibrium could take a very long time compared to the length of a planning horizon. Therefore, in terms of modeling, our work complements existing papers in the literature by considering a transient control problem, which we believe is an important problem in many scenarios. Our choice of modeling the system as a transient system allows us to easily accommodate non-stationary demand arrival process and also common traffic patterns in large cities. It is worth noting here that, since we focus on transient control problem, a lot of techniques for analyzing steady-state systems from the queueing literature that are used in many existing papers noted above are no longer applicable in our setting. As a result, our analysis is different from the existing papers on ride-hailing pricing control.

Secondly, most of the existing papers in ride-hailing literature focus on static controls, whereas we consider both static and dynamic controls to provide insights on the value of dynamic pricing. The only papers we are aware of that focus on dynamic pricing controls are Balseiro et al. (2021), Kanoria and Qian (2019), and Varma et al. (2022). Balseiro et al. (2021) study a Lagragian-relaxation based dynamic pricing control and show that their proposed control is asymptotically optimal in the hub-and-spoke setting where the number of demand regions (more precisely, the number of "spokes") is large. In their model, it is assumed that demand arrival is stationary and servers' travel time between different demand regions is exponentially distributed (i.e., has a memoryless property). Our work complements their modeling approach by studying a model with non-stationary demand arrival process and deterministic travel time between regions. In addition, we focus on a different asymptotic region where the traffic volume and number of servers is large, and develop asymptotically optimal dynamic pricing controls. Kanoria and Qian (2019) develop a joint dynamic control (including admission, matching and pricing decisions), which does not require prior knowledge of the demand arrival rates and is asymptotically optimal in the large supply regime. The key assumption in their model and analysis is the *instantaneous* relocation of cars, i.e., the travel time between regions is always zero. In practice, the fact that it takes a non-negligible time for a driver to move from the origin of the trip to the destination of the trip adds more friction in system dynamic and makes the system slower to respond to dynamic price adjustments. In contrast to their work, we explicitly model such friction and design an effective dynamic pricing control which takes into account the travel time when deciding prices. Varma et al. (2022) consider the joint pricing and matching controls for a two-sided queueing system which is modeled by a bipartite graph where vertices represent customer or server types and the edges represent compatible customer-server pairs. Their model could be applied to ride-hailing systems by treating riders for different origin-destination pairs as different types of customer vertices and drivers at different locations as different types of server vertices. They propose a dynamic joint pricing and matching control which achieves a relative revenue loss in the order of $n^{-2/3}$ in an asymptotic regime when both the supply and demand are scaled up linearly in n. While their asymptotic regime and the performance of their proposed control are similar to ours, our work and their work are different in several ways. First, they focus on the steady-state control problem whereas we consider the transient control problem. Second, the queueing system they consider is an open network; once a server is matched with a customer, both the server and the customer leave the system. In contrast, we consider a closed system where the server remains in the system and becomes available to take new customers once this server finishes the service; this allows us to explicitly capture the spatial and inter-temporal correlations of server availability in our model which is a key feature of the dynamics of ride-hailing systems not captured in Varma et al. (2022).

Aside from the above papers that focus on pricing as the control lever, there is also a growing literature which focuses on other control levers, e.g., matching, driver re-positioning, admission control), e.g., Banerjee et al. (2018), Afeche et al. (2018), Braverman et al. (2019) and Özkan and Ward (2020). All of these papers also focus on steady-state control problem and static controls.

Finally, our work is also related to the stream of literature on real-time dynamic pricing and the literature on revenue management with reusable resources. The former literature seeks to find computationally efficient dynamic pricing controls that can be implemented in real-time (i.e., without having to resort to large-scale re-optimizations) and have provably near-optimal performances; see e.g., Jasin (2014), Chen et al. (2015), Lei et al. (2018), and Lei and Jasin (2020). The second literature studies revenue management problems in which finite resources can be repeatedly used to serve arriving demands after the completion of service time. Recent works have studied such problems in the context of capacity planning (Besbes et al. 2021a), admission control (Levi and Radovanovic 2010, Chen et al. 2017), pricing control (Kim and Randhawa 2018, Besbes et al. 2019, Lei and Jasin 2020), and assortment control (Owen and Simchi-Levi 2018, Feng et al. 2019, Gong et al. 2019, Rusmevichientong et al. 2020). The major difference between our paper and the aforementioned works is that we consider a problem where resources are moving in a *network*, and the location of a resource unit in the network changes upon the completion of service. This creates new challenges that requires new ideas in the design of dynamic pricing controls.

3. Model

In this section, we first discuss the problem setting and modeling assumptions, followed by the analytical framework that we use to evaluate the performance of different pricing controls. At the end of this section, we discuss a simple example that will be used later to illustrate the main ideas of our proposed pricing controls.

3.1. Problem setting and assumptions

Consider a firm that provides a transportation service in a city. We approximate the traffic flow in the city by a network that consists of N nodes indexed by $i = 1, \ldots, N$ and directed arcs between the nodes. Each node i corresponds to a region with a reasonable size so that there is a non-negligible number of trips *within* each region during the horizon and the travel times for these trips are non-negligible. Each arc $i \to j$ corresponds to the trip from region i to region j. We assume that the length of the horizon is one day, and we divide the horizon into T decision periods, each of which is short enough so that, for each period t and any arc $i \to j$, there is at most one potential new customer. At the beginning of each period, the firm needs to quote a price $p_{t,ij} \in \mathcal{P}_{t,ij}$, which is what a customer arriving in period t has to pay to get the service (the trip) from $i \to j$. We denote by $\mathcal{P}_{t,ij} := [\underline{p}_{t,ij}, \bar{p}_{t,ij}] \subset \mathbb{R}_{++}$ the set of feasible prices for trip $i \to j$ in period t. The actual number of $i \to j$ trips induced by $p_{t,ij}$ is denoted by a binary random variable $D_{t,ij}(p_{t,ij}) \in \{0, 1\}$. For each arc $i \to j$ and each period t, we define the *demand function* $\lambda_{t,ij} : \mathcal{P}_{t,ij} \to [0, 1]$ as $\lambda_{t,ij}(p_{t,ij}) := \mathbf{E}[D_{t,ij}(p_{t,ij})]$. Throughout this paper, we will also call $\lambda_{t,ij}(\cdot)$ the *expected demand* or *demand rate* for trip $i \to j$ in period t. Note that we allow the demand function to be non-stationary to capture the predictable variability of average demand rate throughout the day. We also define $\delta_{t,ij}(p_{t,ij}) := D_{t,ij}(p_{t,ij}) - \lambda_{t,ij}(p_{t,ij})$ to capture the unpredictable variability (i.e., the stochasticity) in demand. We make the following assumption for the demand functions.

A1. For any t and any arc $i \to j$, the demand function $\lambda_{t,ij} : \mathcal{P}_{t,ij} \to [0, 1]$ is strictly decreasing and twice continuously differentiable.

The above assumption is standard in the pricing literature and is satisfied by many commonly used demand functions such as linear, logit and exponential. For simplicity, we assume that $\bar{p}_{t,ij}$ is a sufficiently high price that completely turns off the demand for trip $i \to j$ in period t, i.e., $\lambda_{t,ij}(\bar{p}_{t,ij}) = 0$. (Note that although the theoretical turn-off price can be infinite, e.g., when $\lambda_{t,ij}(p) = a \cdot e^{-p}$, since real-world price is never infinite, we can always pick a sufficiently high $\bar{p}_{t,ij} < \infty$ such that the expected demand at $\bar{p}_{t,ij}$ is practically negligible.)

Since $\lambda_{t,ij}(\cdot)$ is *strictly* decreasing, it has an inverse function which we denote as $p_{t,ij} : [0, 1] \to \mathcal{P}_{t,ij}$. The expected revenue from trip $i \to j$ under price $p_{t,ij}$ in period t is equal to $\mathbf{E}[p_{t,ij} D_{t,ij}(p_{t,ij})] =$

$p_{t,ij}\lambda_{t,ij}(p_{t,ij})$. Since $\lambda_{t,ij}$ is invertible, we can also define the expected revenue as a function of demand rate. Specifically, we define $r_{t,ij}(\lambda):[0,1]\to\mathbb{R}_+$ as $r_{t,ij}(\lambda):=\lambda p_{t,ij}(\lambda)$, and make the following assumption on the revenue functions.

A2. For any t and any arc $i\to j$, the revenue function $r_{t,ij}(\lambda):[0,1]\to\mathbb{R}_+$ is strongly concave.

The above assumption is commonly made in the revenue management literature (e.g., Gallego and van Ryzin 1994). Note that although the expected revenue may not be concave in price, it is known to be concave in demand rate for most commonly used demand functions including linear, exponential, and logit. Finally, we make the following regularity assumptions.

A3. There exist positive constants Ψ_1,Ψ_2,Ψ_3 such that: (i) $\lambda'_{t,ij}(p_{t,ij})\in[-\Psi_2,-\Psi_1]$ for all $i,j,t,p_{t,ij}$, and (ii) $\max\{|r'_{t,ij}(\lambda_{t,ij})|,|r''_{t,ij}(\lambda_{t,ij})|\}\leq\Psi_3$ for all $\lambda_{t,ij}$.

Note that **A3** is fairly innocuous. Since $\lambda_{t,ij}$ is strictly decreasing on a compact domain, the derivative of $\lambda_{t,ij}$ is strictly negative and has both finite upper and lower bounds, so **A3**(i) holds. Since $\lambda_{t,ij}$ is twice continuously differentiable, so does $r_{t,ij}$. Consequently, **A3** (ii) holds since $r_{t,ij}$ has a compact domain. Hence, $r_{t,ij}$ is bounded, and we will use $r_{\max}:=\max_{t,i,j,\lambda\in[0,1]}r_{t,ij}(\lambda)$ to denote the maximum unconstrained revenue.

On the supply side, we assume that the firm owns a fixed pool of homogeneous servers (e.g., drivers or a fleet of taxis) to serve the demand. The servers are initially distributed across different regions in the city, and they move around throughout the horizon to serve demands. Specifically, to fulfill a demand in period t for trip $i\to j$, the firm needs to use one server that is currently present in region i in period t to take this customer to region j. From a practical perspective, the restriction of only allowing a server in the same region to fulfill the demand originating in that region is motivated by the fact that customers are delay-sensitive and may turn to other means of transportation if the waiting time is too long. From the modeling perspective, allowing servers from other regions to fulfill demands originating at different regions requires an additional layer of assignment decision, which is beyond the scope of this paper.

We assume that it takes exactly $\tau_{ij}\geq 0$ periods for a server to travel from region i to j, during which the server is "busy" and cannot take a new customer. In other words, a server who picks up the customer in region i in period t will be busy from periods t to $t+\tau_{ij}-1$, and then becomes available again in period $t+\tau_{ij}$ in region j. We denote by $\underline{\tau}:=\min_{i,j}\tau_{ij}$ the minimum travel time across all arcs, and assume that $\underline{\tau}>0$ to capture the non-negligible travel time. The deterministic travel time assumption is different from the existing literature (e.g., Balseiro et al. 2021, Kanoria and Qian 2019, and Varma et al. 2022) where travel time either is assumed to be zero or follow an exponential distribution. We choose this modeling approach to capture the heterogeneous travel time across different trips; this choice is also motivated by the fact that travel time may not be "memoryless" (as is the case with exponential distribution). In practice, travel time can be highly uncertain and is not always easy to accurately approximate using either a deterministic or an exponential random variable. Addressing the problem of dynamic pricing with a general service time distribution is an interesting yet challenging problem. We leave this for future research.

We use the vector $C_t = (C_{t,1}, \ldots, C_{t,N})^\top$ to denote the number of servers that are available for service at different regions at the beginning of period t. For convenience, we assume that there are at least N servers available in each region at the beginning of the horizon, i.e., $C_{1,i} \geq N$ for any region i. (This is not a restrictive assumption since we focus on asymptotic analysis. Moreover, since major taxi companies in large cities typically have a large fleet of taxis and we take a modeling approach where each region has a reasonable size, there will be sufficient number of taxis in each region.) As mentioned previously, throughout the horizon, C_t evolves stochastically because servers move around from one region to another region as they pick up and drop off customers. Similar to Balseiro et al. (2021), Kanoria and Qian (2019), and Varma et al. (2022), we assume no strategic behaviors on both the demand and supply side: Demand responds directly to the current period quoted prices without either referencing to the historic price trajectory or anticipating a future price trajectory, and servers do not relocate themselves across the network for their own interest.

3.2. Problem formulation, approximations, and performance metric

The firm's objective is to find an *admissible pricing control* to maximize its cumulative expected revenue during the horizon. Mathematically, an admissible pricing control is a sequence of functions which determines the prices for each arc in each period, i.e., $\pi = \{\pi_t : \mathcal{F}_t \to \mathbb{R}_+^{N^2}\}_{t=1}^T$ where $\mathcal{F}_t$ denotes all the information that is available at the beginning of period t. Let Π denote the class of admissible pricing controls. The firm's dynamic pricing problem can be formulated as follows:

$$(\mathbf{SDP}) \qquad \mathcal{J}^\star := \max_{\pi\in\Pi} \quad \mathbf{E}^\pi \left[\sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N p^\pi_{t,ij} D_{t,ij}(p^\pi_{t,ij}) \right] \tag{1}$$

$$\text{s.t.} \quad C_i - \sum_{j=1}^N \sum_{s=1}^t D_{s,ij}(p^\pi_{s,ij}) + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji})^+} D_{s,ji}(p^\pi_{s,ji}) \geq 0, \ \forall i, t, \tag{2}$$

$$p^\pi_{t,ij} \in \mathcal{P}_{t,ij} \ \forall i, j, t, \tag{3}$$

where $p^\pi_{t,ij}$ is the price under policy π, the constraints must hold almost surely and the expectation in (1) is taken with respect to the stochastic processes induced by the control π. Constraint (2) is a set of flow-balance constraints that requires, in each period t, the number of servers in each node cannot be negative. To better understand the expression in (2), consider the number of available servers at node i at the beginning of period t. By the end of period $t-1$, there has been a total $\sum_{j=1}^{N}\sum_{s=1}^{t-1} D_{s,ij}(p^{\pi}_{s,ij})$ rides that originate from node i to some destination node j. At the same time, by the end of period $t-1$, there has also been a total of $\sum_{j=1}^{N}\sum_{s=1}^{(t-\tau_{ji})^+} D_{s,ji}(p^{\pi}_{s,ji})$ rides that have arrived at node i from some origin node j. Therefore, the number of available servers at the beginning of period t in node i can be computed as $C^{\pi}_{t,i} := C_i - \sum_{j=1}^{N}\sum_{s=1}^{t-1} D_{s,ij}(p^{\pi}_{s,ij}) + \sum_{j=1}^{N}\sum_{s=1}^{(t-\tau_{ji})^+} D_{s,ji}(p^{\pi}_{s,ji})$. The pricing decision for the rides originating in node i at period t must satisfy $C^{\pi}_{i,t} \geq \sum_{j=1}^{N} D_{t,ij}(p^{\pi}_{t,ij})$, which is exactly (2).

While in theory the stochastic optimization problem defined above can be solved by dynamic programming, in practice, it is computationally intractable due to curse of dimensionality (e.g., its states include, for each node, the number of idle servers and, for each arc, the number of servers that will arrive in the destination in every single future periods). Therefore, in this paper, we focus on developing provably near-optimal heuristic controls. Since the optimal expected revenue is hard to compute, in order to analytically and numerically evaluate the effectiveness of any proposed heuristic pricing control, we develop an upper bound of the optimal expected revenue. To this end, we first introduce a class of parameterized *deterministic convex optimizations.*

For any $\zeta \in \mathbb{R}_+$, we define

$$(\mathbf{DCP}(\zeta)) \quad \mathcal{J}(\zeta) := \max_{\{\lambda_{t,ij}\}_{t,i,j}} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} r_{t,ij}(\lambda_{t,ij}) \tag{4}$$

$$\text{s.t.} \quad C_i - \sum_{j=1}^{N}\sum_{s=1}^{t} \lambda_{s,ij} + \sum_{j=1}^{N}\sum_{s=1}^{(t-\tau_{ji})^+} \lambda_{s,ji} \geq 0, \forall i, t \tag{5}$$

$$\zeta \leq \lambda_{t,ij} \leq 1-\zeta, \forall t, i, j \tag{6}$$

where, by convention, we set $\mathcal{J}(\zeta) = -\infty$ whenever $\mathbf{DCP}(\zeta)$ is not feasible. Note that in the above we use the expected demand $\lambda_{t,ij}$ as the immediate decision variable instead of the price $p_{t,ij}$. This is possible since the demand function $\lambda_{t,ij}(\cdot)$ is invertible. Compared to the classic deterministic relaxation in many revenue management literature (e.g., Gallego and van Ryzin 1994), $\mathbf{DCP}(\zeta)$ is a perturbed optimization problem parameterized by ζ. The purpose of ζ is to make sure that the optimal solution of $\mathbf{DCP}(\zeta)$ lies in a proper interior of $[0,1]$, which is important for the analysis of our heuristic controls; we provide more detailed discussions on this technical issue in Remark 1.

Let $\{\lambda^{\zeta,D}_{t,ij}\}_{t,i,j}$ denote the optimal solution to $\mathbf{DCP}(\zeta)$, and by $\{p^{\zeta,D}_{t,ij} := p_{t,ij}(\lambda^{\zeta,D}_{t,ij})\}_{t,i,j}$ the corresponding price solution. Note that when $\zeta = 0$, $\mathbf{DCP}(0)$ becomes a *deterministic approximation* of **SDP** where we simply replace all the random variables in **SDP** with their expected values. The following result shows that $\mathbf{DCP}(0)$ is an upper bound of **SDP**.

LEMMA 1. *$\mathcal{J}(0) \geq \mathcal{J}^{\star}$.*

For any admissible control π that satisfies (2) and (3), let R^π denote the cumulative revenue during the horizon. Lemma 1 motivates us to define the following performance metric

$$\mathcal{L}^\pi = \frac{\mathcal{J}(0) - \mathbf{E}^\pi[R^\pi]}{T}, \tag{7}$$

which can be interpreted as the average per period revenue loss (henceforth, *loss* for short).

Remark 1. The class of deterministic convex optimizations $\{\mathbf{DCP}(\zeta)\}_\zeta$ introduced above is motivated by the framework employed in the price-based revenue management literature (e.g., Gallego and van Ryzin 1994). In this literature, however, it is sufficient to introduce a *single* deterministic approximation (similar to $\mathbf{DCP}(0)$) to simultaneously achieve two objectives: First, its optimal objective value is an upper bound of the optimal revenue, and can be used to define a proper performance metric (similar to *loss* defined in our setting); second, its optimal solution can be used to develop effective heuristic controls. In fact, one popular approach for developing effective heuristic pricing control in this literature is as follows: Use the optimal solution of the deterministic approximation as a *static baseline control*, and then adaptively perturb the baseline control to account for stochastic variability of demand. For this approach to work, it typically requires the baseline control to satisfy certain "interior" condition so that there is some room to perturb the baseline control either up or down (see, for example, Jasin 2014). While such condition is naturally satisfied in the standard price-based revenue management problem discussed in the classic literature, unfortunately, it may not hold in our problem even for reasonable problem instances. Specifically, due to the non-stationary of demand and the complex movement of servers in a network, for some arc $i \to j$ in some period t, the solution to $\mathbf{DCP}(0)$ may coincide with the boundary of the domain, say 0 (resp. 1); in this case, the seller can no longer decrease (resp. increase) the demand rate when needed. (Indeed, in one of the numerical experiments, the expected demand of some arc at the deterministic optimal solution is zero; see Figure 2a.) This is why in our problem, we need to introduce the "cushion parameter" ζ and consider a class of deterministic convex optimizations in order to develop our heuristics: When $\zeta > 0$ and $\mathbf{DCP}(\zeta)$ is feasible, it is guaranteed that the firm can adjust the demand rate above or below $\lambda_{t,ij}^{\zeta,D}$ by ζ without violating the boundary of the domain. While increasing ζ provides more flexibility in adjusting the baseline control, it comes at a cost of a more "biased" deterministic approximation of the original stochastic problem. Our proposed heuristics in Sections 4 and 5 will tune the cushion parameter to balance this trade-off.

3.3. The asymptotic regime

The performance metric $\mathcal{L}^\pi$ is hard to characterize analytically, which motivates us to resort to an asymptotic approach. The asymptotic regime we consider in this paper is one where both the total number of potential customers and the total number of servers are large. This is a reasonable regime to consider since in big cities such as in Singapore, there are on average about 563,000 daily completed trips and about 15,000 taxis (Yong 2022). To operationalize this asymptotic setting, we consider a sequence of problem instances whose model primitives are scaled up by a sequence of scaling factors: For any scaling factor $n \in \mathbb{Z}_+$, in the corresponding n^{th} problem, we scale the following model primitives while keeping everything else the same

$$C^n_{1,i} = nC_{1,i}\,; \quad T^n = nT\,; \quad \tau^n_{ij} = n\tau_{ij},; \quad \lambda^n_{t,ij}(p) = \lambda_{\lceil t/n \rceil, ij}(p), \ \ \forall t,i,j,p; \tag{8}$$

(consequently, $p^n_{t,ij}(\lambda) = p_{\lceil t/n\rceil,ij}(\lambda)$, and $\underline{\tau}^n = n\underline{\tau}$). Note that, in this sequence of scaled problems, the total demand volume and the total number of servers are scaled up at the same rate. In particular, since we assume at most one arrival per period per arc, the scaling in the length of the horizon and the service time reflects a change in time scale rather than in the actual length of the horizon and service time. In other words, a scaled system has the same actual length of horizon and service time compared to the unscaled system, but the length of one period is smaller than that in the unscaled system. Consequently, the scaling in the service time should be interpreted as increasing the demand arrival intensity as well as the frequency at which firm makes the pricing decisions. The scaling in the demand function essentially assumes that the actual demand arrival is piece-wise stationary, which is a commonly used assumption in the transportation literature to estimate demand (Yao et al. 2018).

For the scaled system with factor n, we also define the corresponding scaled stochastic problem $\mathbf{SDP}(n)$ and deterministic approximations $\mathbf{DCP}(n,\zeta)$. Let $\mathcal{J}(n,\zeta)$ denote the optimal objective value of $\mathbf{DCP}(n,\zeta)$, and $\lambda^{n,\zeta,D}_{t,ij}$ denote its optimal solution. The following holds for $\mathbf{DCP}(n,\zeta)$.

Lemma 2. *There exist constants $\bar{\zeta} > 0$ and $\Psi_0 > 0$ independent of n and ζ such that for all $\zeta < \bar{\zeta}$, $\mathcal{J}(n,0) - \mathcal{J}(n,\zeta) \le \Psi_0 T^n \zeta$.*

Let $\mathbf{E}^\pi[R^\pi(n)]$ denote the expected revenue under control π in the n^{th} scaled problem. We are interested in characterizing the *asymptotic order* of the loss of π as a function of n:

$$\mathcal{L}^\pi(n) := \frac{\mathcal{J}(n,0) - \mathbf{E}^\pi[R^\pi(n)]}{T^n}. \tag{9}$$

In the remainder of the paper, for notational simplicity, we will simply suppress the superscript n whenever there is no confusion.

3.4. A simple example

We now provide an example of the model we have introduced. Consider a planning horizon of $T = 30$ decision periods for a network of $N = 3$ nodes. Suppose that, at the beginning of the horizon, there are 10 servers in each node, i.e., $C_1 = (10, 10, 10)^\top$. For simplicity, we assume there are no demand for trips whose origin and destination are the same, i.e., $\lambda_{t,ii}(p) = 0$ for all i and t. Suppose that the travel time between any two nodes $i \neq j$ is $\tau_{ij} = 10$ periods, and its expected demand equals $\lambda_{t,ij}(p) = 1 - 4p$ when $t \leq 10$, and equals $\lambda_{t,ij}(p) = 1 - \frac{p}{2}$ when $t > 10$. One can view the last 20 periods as *peak* periods where the demand is less elastic and riders are less price sensitive compared to the first 10 *off-peak* periods. We can formulate **DCP** as follows (for illustration purposes, we work with the case by letting $\zeta = 0$):

$$
\begin{aligned}
\mathcal{J}(0) := \max_{\{\lambda_{t,ij}\}_{t,i,j}} \quad & \sum_{t=1}^{10} \sum_{i=1}^{3} \sum_{j:j\neq i} \frac{(1-\lambda_{t,ij})\lambda_{t,ij}}{4} + \sum_{t=11}^{20} \sum_{i=1}^{3} \sum_{j:j\neq i} 2(1-\lambda_{t,ij})\lambda_{t,ij} \\
\text{s.t.} \quad & 10 - \sum_{j:j\neq i} \sum_{s=1}^{t} \lambda_{s,ij} + \sum_{j:j\neq i} \sum_{s=1}^{t-10} \lambda_{s,ji} \geq 0, \forall i, t \\
& 0 \leq \lambda_{t,ij} \leq 1, \forall t, i, j
\end{aligned}
$$

It can be easily verified that $\{\lambda^{0,D}_{t,ij} = \frac{1}{2}\}_{t,i,j}$ (in fact, $\{\lambda^{0,D}_{t,ij} = \frac{1}{2}\}_{t,i,j}$ is also the unconstrained optimal solution); thus, the deterministic optimal prices for the trips between nodes $i \neq j$ are given by:

$$
p^{0,D}_{t,ij} = \begin{cases} \frac{1}{8}, \text{ for all } t \leq 10 \\ 1, \text{ for all } t > 10 \end{cases}. \tag{10}
$$

It is worth noting that while in the optimal solution of **DCP** the optimal demand rate is the same during the peak and off-peak periods, the revenue rate is much higher during the peak periods due to a reduction in riders' price sensitivity. From the firm's perspective, the peak periods are much more profitable than the off-peak periods. We will use this example to illustrate the main idea of the heuristics we propose in later sections.

4. Static Pricing

In this section, we focus on the class of static pricing controls where the firm can only choose the prices before the planning season starts and cannot adaptively adjust the prices, and investigate the performance of this class of pricing controls.

4.1. A lower bound of loss under static pricing

In this subsection, we will construct a problem instance and show that no static pricing control can achieve an asymptotic loss smaller than $n^{-1/2}$.

Consider an unscaled problem instance of a network with $N=2$ nodes and a horizon of $T=10$ periods. On the demand side, $\lambda_{ii}(p)=0$ for $i=1,2$, and $\lambda_{ij}(p)=1-p$ for $i,j\in\{1,2\}$ and $j\neq i$. In other words, there are no trips within the same region but only trips across the regions; moreover, the demand pattern is stationary across the horizon. The travel time between the two regions is two periods, i.e., $\tau_{12}=\tau_{21}=\tau:=2$. On the supply side, $C_1=C_2=1$. The following result holds.

Theorem 1. *In the problem instance described above, for any static price control π, there exists some constant $\tilde{\Psi}_0>0$ that is independent of the policy π such that $\mathcal{L}^{\pi}(n)\geq\tilde{\Psi}_0\cdot n^{-1/2}$.*

The intuition of the proof of the lower bound above is as follows: We first lower bound the total revenue loss of any control by its expected revenue loss during the first τ periods. Since no server will arrive at its destination during the first τ periods, the number of servers at every node cannot increase during the same time periods. Therefore, we can further link the pricing problem during the first τ periods to the classic network revenue management problem with finite capacity, which is known to have a revenue loss of at least $\Omega(n^{-1/2})$. We would like to note that while Theorem 1 tells us that the gap between any static policy and the deterministic upper bound is at least $\Omega(n^{-1/2})$, it does not directly imply that the gap between any static policy and the optimal policy is at least $\Omega(n^{-1/2})$. The latter can be established if there is a policy whose loss is sub-$n^{-1/2}$, which is the case for our dynamic policy; we will come back to this point later in Section 5.

4.2. A static pricing control

We now discuss a static pricing control that guarantees a loss close to $n^{-1/2}$. To motivate the ideas behind our control, recall that, as mentioned in the previous section, one effective approach employed in the price-based revenue management literature is to use the optimal solution to the deterministic approximation as a static pricing control as long as the firm still has enough resource to accommodate the demand (Gallego and van Ryzin 1994). While this idea seems reasonable, this approach may not be directly applicable to our context. To illustrate this, consider the example in Section 3.4. To apply the classic static pricing control approach for price-based revenue management in our example, the firm should set $p_{t,ij}=p^{0,D}_{t,ij}$, i.e., $p_{t,ij}=\frac{1}{8}$ for all trips in the first 10 periods, and then apply $p_{t,ij}=1$ in the remaining 20 periods. This means that, on average, in any period, the number of servers going out of any node i equals $\sum_{j\neq i}\lambda^{0,D}_{t,ij}=1$. Consider a hypothetical scenario where there is no demand variability. Recall that the travel time is 10 periods; so each node would not get any servers from other nodes until period 11 and would experience a net reduction of one server in the first 10 periods; afterwards, node i would have an inflow of $\frac{1}{2}$ server from each of the other two nodes, which would cancel out the outflow of one server from node i. Since the initial number of servers in each node equals 10, each node would have zero server available at the end of periods 11 to 30. (Mathematically, this corresponds to flow-balance constraints in **DCP** that are binding at its optimal solution.) The key implication of this hypothetical scenario is that any demand variability could result in *demand blockage* during the last 20 periods. Indeed, consider a possible scenario where the total outflow from node 1 during the first 10 periods equals its average (i.e., 10 servers), but the actual trips from other nodes to node 1 in period 1 are less than average, e.g., $D_{1,21} = D_{1,31} = 0$. Then, even though the total demand was less than average, node 1 would have no server available to pick up riders in period 11. From the firm's perspective, this is an undesirable situation since the firm has to deny riders during peak hours simply because it has local supply shortage in its network.

As the above example illustrates, the root cause of *demand blockage* is the underlying variability of demand, which cannot be directly addressed unless the firm leverage adaptive adjustment of prices. However, it is still possible to use static pricing controls to *mitigate* the impact of demand blockage. One intuitive fix for this blockage problem is to maintain buffer servers in each node i to deal with demand variability. Specifically, if the firm reduces the average outgoing flows by slightly increasing the static prices for all the outgoing flows in node i, then it can create some buffer servers in node i. However, reducing the outflow in node i in the current period implies that the inflow in other nodes in the future are reduced. Hence, we need to adjust the static prices in a *balanced* way so that on average all nodes can maintain a sufficient level of buffer servers.

We propose the following idea: Set the static price $p_{t,ij}$ slightly higher than $p^{\zeta,D}_{t,ij}$ so that the average demand for each arc $i \to j$ in all periods is reduced by the same quantity ϵ, i.e., $\lambda_{t,ij}(p_{t,ij}) = \lambda^{\zeta,D}_{t,ij} - \epsilon$. To illustrate why this would help build up buffer in a *balanced* way, suppose there is no demand variability in the system and it takes the same τ periods for trips between any two nodes. Now, consider any node i. Since the inflows take τ periods to reach node i, their impact on the number of available servers in node i will not occur until after period $\tau + 1$; meanwhile, the firm is reducing each of the N outflows by ϵ per period for τ periods; thus, the firm can build up extra buffer of $N\epsilon\tau$ at the end of period τ. Afterwards, the reduction in all N outflows perfectly compensates the reduction in all N inflows. As the flow adjustments balance out, node i would keep $N\epsilon\tau$ buffer servers in the remainder of the planning horizon. Let $\textsc{Proj}_{t,ij}(.)$ denote the projection to $\mathcal{P}_{t,ij}$. We formally introduce our proposed *static price control* (SPC) below.

Static Price Control

Input: Tuning parameters ϵ, ζ

Step 1: Solve $\mathbf{DCP}(\zeta)$ to obtain $\{\lambda_{t,ij}^{\zeta,D}\}_{t,ij}$.

Step 2: For $t=1$ to T, do for each i:

a. If $C_{t,i} \geq N$, set $p_{t,ij}^{SPC} = \textsc{Proj}_{t,ij}(p_{t,ij}(\lambda_{t,ij}^{\zeta,D} - \epsilon))$ for all j;

b. Otherwise, set $p_{t,ij}^{SPC} = \bar{p}_{t,ij}$.

The following result characterizes the performance of SPC.

Theorem 2. *Set* $\epsilon = \zeta = (\underline{T}^n)^{-1}(1+\sqrt{16T^n \log(\underline{T}^n)})$. *There exists some constant* $\tilde{\Psi}_1 > 0$ *independent of* n *such that*

$$\mathcal{L}^{SPC}(n) \leq \tilde{\Psi}_1 \cdot \sqrt{\frac{\log(1+n)}{n}}.$$

Comparing Theorem 2 with Theorem 1, we can see that the performance of SPC matches the theoretical lower bound on any static pricing control by at most a multiplicative logarithmic factor.

4.3. Outline of the proof of Theorem 2

Let $\pi = SPC$. The proof of Theorem 2 is divided into several steps: We first argue that, under the prescribed choice of ϵ and ζ, the bound of loss of SPC holds when the scaling parameter n is "small". When n is large, we define an event on which (i) the number of servers in each node is always positive throughout the horizon, and (ii) the desirable demand rate after the adjustment lies in the interior of the feasible demand range. We then proceed to bound the loss of SPC by conditioning on the event defined above.

Analysis of small n

By definition of ϵ and (8),

$$\epsilon = \frac{1}{n\underline{T}} + \frac{4\sqrt{T}}{\underline{T}}\sqrt{\frac{\log(\underline{T}^n)}{n}} \leq \frac{\sqrt{T}}{\underline{T}}\sqrt{\frac{1}{n}} + \frac{4\sqrt{T}}{\underline{T}}\sqrt{\frac{\log(n\underline{T})}{n}} < \frac{5\sqrt{T}}{\underline{T}}\sqrt{\frac{1+\log(n\underline{T})}{n}}. \tag{11}$$

Note that the right hand side of (11) is decreasing in $n \in \mathbb{Z}_{++}$ and converges to 0.

Let $\Omega := \max\{n \in \mathbb{Z}_{++} : \sqrt{1+\log(n\underline{T})/n} > (5\sqrt{T})^{-1}\bar{\zeta}\underline{T}\}$ (if the right hand side is an empty set, let $\Omega := 0$). Then, for all $n \in \mathbb{Z}_{++}$ such that $n \leq \Omega$,

$$\mathcal{L}^{\pi}(n) \leq \frac{\mathcal{J}(n,0)}{T^n} \leq \frac{N^2 T^n r_{\max}}{T^n} = N^2 r_{\max} \leq M_1 \frac{\sqrt{T}}{\underline{T}}\sqrt{\frac{1+\log(n\underline{T})}{n}}$$

where the first inequality holds by $\mathbf{E}^{\pi}[R^{\pi}(n)] \geq 0$, the second inequality follows by the definition of $r_{\max}$, the last inequality holds by defining $M_1 := 5N^2 r_{\max}/\bar{\zeta}$ which is independent of $T, \underline{T}$, and n, and the fact that $5\sqrt{T}(\underline{T}\bar{\zeta})^{-1}\sqrt{1+\log(n\underline{T})/n} > 5\sqrt{T}(\underline{T}\bar{\zeta})^{-1}\sqrt{1+\log(\Omega\underline{T})/\Omega} > 1$ for all $n \leq \Omega$.

Analysis of large n

When $n > \Omega$, by the definition of Ω and (11), the following condition holds,

$$\mathbf{C1}: \quad \bar{\zeta} > \epsilon = \zeta,$$

which implies $\mathbf{DCP}(n,\zeta)$ is feasible and $\lambda_{t,ij}^{n,\zeta,D}$ is well-defined. Note that for all $i \to j, t$, we have $\lambda_{t,ij}^{n,\zeta,D} - \epsilon = \lambda_{t,ij}^{n,\zeta,D} - \zeta \in (0,1)$, so $p_{t,ij}^{\pi} = p_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D} - \epsilon)$. Define the set

$$\mathcal{S}_{ij} := \left\{ \max_{1 \le t \le T^n} \left| \sum_{s=1}^{t} \delta_{s,ij}^{n} \right| < \frac{(\epsilon \underline{T}^n - 1)}{2} \right\}$$

and $\mathcal{S} := \cap_{i,j} \mathcal{S}_{ij}$. It can be shown that, condition **C1** implies that supply at all the nodes will not run out throughout the horizon with a high probability. We state a lemma.

Lemma 3. *If **C1** holds, then for all sample paths on $\mathcal{S}$, the following condition holds for all t,*

$$\mathbb{H}_t: \quad C_{t,i}^{n} \ge N \text{ for all } i, \text{ and } p_{t,ij}^{\pi} = p_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D} - \epsilon) \text{ for all } i,j.$$

Moreover, $\mathbf{P}(\mathcal{S}) \ge 1 - 2N^2/(n\underline{T})$.

Using Lemma 3, the loss under SPC when $n > \Omega$ can be bounded as follows:

$$\begin{aligned}
\mathcal{J}(n,\zeta) - \mathbf{E}^{\pi}\left[R^{\pi}(n)\right] &= \mathcal{J}(n,\zeta) - \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{E}^{\pi}\left[\mathbf{E}^{\pi}\left[p_{t,ij}^{\pi} D_{t,ij}^{n}(p_{t,ij}^{\pi}) \middle| \mathcal{F}_t\right]\right] \\
&= \mathcal{J}(n,\zeta) - \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{E}^{\pi}\left[r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi}))\right] \\
&= \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N} \left\{ \mathbf{E}^{\pi}\left[r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi})) \middle| \mathcal{S}\right] \mathbf{P}(\mathcal{S}) \right. \\
&\qquad \left. + \mathbf{E}^{\pi}\left[r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi})) \middle| \mathcal{S}^c\right] \mathbf{P}(\mathcal{S}^c) \right\} \\
&\le \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N} \left\{ \mathbf{E}^{\pi}\left[r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi})) \middle| \mathcal{S}\right] \mathbf{P}(\mathcal{S}) \right. \\
&\qquad \left. + r_{\max}\mathbf{P}(\mathcal{S}^c) \right\}
\end{aligned} \tag{12}$$

where the second equality follows by the definition of the revenue function, the third equality follows by law of total expectation, and the inequality follows since $r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi})) \le r_{\max}$. We further bound the first term after the last inequality in (12). Recall that, when $n > \Omega$, $\mathbb{H}_t$ holds on $\mathcal{S}$; so $\lambda_{t,ij}^{n}(p_{t,ij}^{\pi}) = \lambda_{t,ij}^{n,\zeta,D} - \epsilon$ (i.e. the projection operator in Step 1 is always inactive). Thus,

$$\mathbf{E}^{\pi}\left[r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n}(p_{t,ij}^{\pi})) \middle| \mathcal{S}\right] = r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^{n}(\lambda_{t,ij}^{n,\zeta,D} - \epsilon^n) \le \Psi_3 \epsilon, \tag{13}$$

where the inequality follows by **A3**.

Combining (12) and (13), we have

$$\begin{aligned}
\mathcal{L}^\pi(n) &= \frac{\mathcal{J}(n,0)-\mathcal{J}(n,\zeta)+\mathcal{J}(n,\zeta)-\mathbf{E}^\pi[R^\pi(n)]}{T^n} \\
&\leq \Psi_0\zeta + N^2\left(\Psi_3\epsilon\,\mathbf{P}(\mathcal{S}) + r_{\max}\cdot\frac{2N^2}{n\underline{\tau}}\right) \\
&\leq \Psi_0\zeta + N^2\left(\Psi_3\epsilon + r_{\max}\cdot\frac{2N^2}{n\underline{\tau}}\right) \\
&= (\Psi_0+N^2\Psi_3)\epsilon + 2N^4 r_{\max}(n\underline{\tau})^{-1} \leq M_2\frac{\sqrt{T}}{\underline{\tau}}\sqrt{\frac{1+\log(n\underline{\tau})}{n}},
\end{aligned}$$

where the first inequality follows by Lemma 2 and 3, the second inequality follows since $\mathbf{P}(\mathcal{S})\leq 1$, the third inequality follows by (11), the definition of $M_2 = 5(\Psi_0+N^2\Psi_3)+2N^4 r_{\max}$, and the fact that $n^{-1}\leq\sqrt{T(1+\log(n\underline{\tau}))/n}$. Setting $\tilde{\Psi}_1 = \max\{M_1, M_2\}\sqrt{T[1+(1+\log(\underline{\tau}))/\log(2)]}\underline{\tau}^{-1}$ completes the proof. □

Remark 2. We briefly explain the intuition behind the asymptotic order of the loss. First, note that $\lambda_{t,ij}^{\zeta,D}\geq\zeta$ (this is achieved by the perturbed deterministic approximation $\mathbf{DCP(\zeta)}$). Since we choose $\epsilon=\zeta$, it is guaranteed that the actual demand rate is smaller than $\lambda_{t,ij}^{\zeta,D}$ by ϵ. Given this and constraint (5), for each node i, SPC is designed to reserve at least $\sum_{j=1}^N \epsilon\tau_{ji}^n \geq N\cdot\epsilon\underline{\tau}^n$ servers unused in every node to absorb the potential demand variability. On the other hand, the maximum amount of random variability in the consumption of servers in each node is on the same order of the total demand variability across T^n periods. Since the demand per period is a Bernoulli random variable, if we set $\epsilon\cdot\underline{\tau}^n\approx\sqrt{T^n}$, then the servers reserved will not be depleted with high probability. (This explains the definition of $\mathcal{S}_{ij}$; the probability bound follows from a large deviation argument, which we formalize in Lemma EC.1.) Therefore, we need to set $\epsilon=\tilde{\Theta}(\sqrt{T^n}/\underline{\tau}^n)=\tilde{\Theta}((\sqrt{T}/\underline{\tau})\cdot n^{-1/2})$. The loss is on the same order of ϵ since the revenue function is Lipschitz continuous.

5. Dynamic Pricing

In this section, we develop a dynamic pricing control where prices are adjusted adaptively. This added flexibility allows the firm to respond adaptively to the observed variability in the system in real-time to better match the supply and demand on the network over time. In what follows, we first discuss our heuristic control and its performance, and then we provide an outline of the proof of our main result in this section.

5.1. Description of the heuristic

To motivate the idea of our dynamic pricing control, let us first revisit the main reason why we could not achieve a loss bound smaller than $n^{-1/2}$ with static controls. Recall that, to avoid demand blockage, we need to preserve extra servers in each node in order to absorb the impact of stochastic variability in demand for all trips that either originate or terminate at that node. Since prices cannot be adjusted adaptively under static controls, the impact of stochastic variability in demand is proportional to the standard deviations of total demand during the *entire horizon*. This in turn determines the degree of suboptimality of the static prices compared to the deterministically optimal prices. However, if we can adjust prices adaptively, we could potentially reduce the amount of stochastic variability that the buffers need to compensate for, thus, reducing the degree of suboptimality. To achieve this, we propose the following variability correction mechanism: We divide the horizon into small batches (intervals) and the demand rates are adjusted such that the cumulative variability in the previous batch is "corrected" in the current batch.

Before formally introducing the details of this variability correction mechanism, we first revisit the example in Section 3.4 to illustrate the main idea. Suppose we want to adaptively make price adjustment every $b = 10$ periods (i.e., this results in three batches within the planning horizon of $T = 30$ periods), and suppose that, building on the buffer server idea developed in SPC, we set $\epsilon = \frac{1}{10}$, and, for ease of explanation, we set $\zeta = 0$. This implies that, for any period and any trip, if the prices remain unadjusted, i.e., $p^{\pi}_{t,ij} = p_{t,ij}(\lambda^{0,D}_{t,ij} - \epsilon)$, then the average demand equals $\lambda_{t,ij}(p^{\pi}_{t,ij}) = \lambda^{0,D}_{t,ij} - \epsilon = \frac{1}{2} - \frac{1}{10} = \frac{2}{5}$. Let us focus on a particular arc $1 \to 2$, and consider the hypothetical scenario where there is no demand variability as our baseline scenario: By the end of period $\tau = 10$, the cumulative flow from node 1 to 2 equals $10 \times \frac{2}{5} = 4$ servers. However, due to stochastic variability, it is possible that the actual cumulative flow in the first 10 periods is more than 4, say 6 servers. To compensate for this extra flow of 2 servers, the idea is to reduce the flow rate in the next batch (i.e., periods 11 to 20) uniformly: Instead of setting prices to induce the target demand rate of $\lambda^{0,D}_{t,12} - \epsilon = \frac{2}{5}$ for $11 \le t \le 20$, we distribute the reduction of 2 servers uniformly across all 10 periods by selecting prices to induce a rate of $\lambda^{0,D}_{t,12} - \epsilon - \frac{2}{10} = \frac{1}{5}$. Thus, by the end of second batch (i.e., period 20), the stochastic variability of the flow $1 \to 2$ incurred in the first batch will be fully compensated and, in contrast to the static controls, will not have any lasting impact on server level during the remaining 10 periods. Of course, extra stochastic variability would be incurred in the second batch; we compensate this in a similar fashion in the third batch.

While the above example provides a high-level intuition on how our variability correction mechanism works, its exact details is more nuanced, which we explain below starting with the definition of batches. Let $\mathcal{T}^k_{ij}$ be the k^{th} batch for arc $i \to j$, which contains several consecutive time periods.

We define the *cumulative (stochastic) demand variability* (i.e., the difference between the realized demand and the demand rate) during the k^{th} batch for arc $i \to j$ as follows:

$$\bar{\delta}_{ij}^k := \sum_{t \in \mathcal{T}_{ij}^k} \delta_{t,ij} = \sum_{t \in \mathcal{T}_{ij}^k} (D_{t,ij}(\lambda_{t,ij}) - \lambda_{t,ij}),$$

and we also define $\kappa_{ij}(t)$ to be the index of the batch which period t belongs to (i.e. $t \in \mathcal{T}_{ij}^{\kappa_{ij}(t)}$). (To be precise, both $\delta_{t,ij}$ and $\lambda_{t,ij}$ depends on the prescribed policy π; yet we ignore such dependency in the notation for brevity.) Let $\chi_{ij}(k) = \arg\max_s \{s \in \mathcal{T}_{ij}^{\kappa_{ij}(t)}\}$ denote the index for the last period in batch k for arc $i \to j$. We define batches sequentially in the following way. First, define $\mathcal{T}_{ij}^0 = \emptyset$ (which implies that $\bar{\delta}_{ij}^0 = 0$) and $\chi_{ij}(0) = 0$. At the end of batch $k \geq 0$, define the end of the next batch $k+1$ as

$$\chi_{ij}^{k+1} := \min \left\{ \chi_{ij}(k) < s \leq T \ : \ \sum_{v=\chi_{ij}(k)+1}^{s} \lambda_{v,ij}^{\zeta,D}(1 - \lambda_{v,ij}^{\zeta,D}) \geq b \right\} \tag{14}$$

where b is a parameter to be chosen which we discuss in more detail below.

In order to understand the intuition behind (14), let us assume that $\bar{\delta}_{ij}^k \geq 0$. In this case, since the total realized demands on arc $i \to j$ is larger than the total demand rates, it would be reasonable to decrease the demand rate in the next batch $k+1$. In particular, we want the total amount of demand rate adjustments in the next batch $k+1$ to be exactly equal to $\bar{\delta}_{ij}^k$. If such correction is possible, then, by the end of batch $k+1$, the demand variability induced during batch k will be completely canceled out and have no impact on the server level in either node i or j in the future. Note that, such correction is only effective if there is enough room for price adjustment given the baseline price control. In particular, since the baseline static demand rate is $\lambda_{t,ij}^{\zeta,D}$ and we plan to *decrease* the demand rates in batch $k+1$ (recall that we assume $\bar{\delta}_{ij}^k \geq 0$), the room of adjustment for each period t in batch $k+1$ is in the order of $\lambda_{t,ij}^{\zeta,D}$. Conversely, in the other case where the total demand variability in batch k is negative (i.e. $\bar{\delta}_{ij}^k < 0$), we will make adjustment by increasing the demand rates in batch $k+1$, so the room of adjustment for each period t in batch $k+1$ is in the order of $(1 - \lambda_{t,ij}^{\zeta,D})$. Thus, regardless of the sign of $\bar{\delta}_{ij}^k$, the room of adjustment for period t in batch $k+1$ is *at least* $\lambda_{t,ij}^{\zeta,D}(1 - \lambda_{t,ij}^{\zeta,D})$. Therefore, we choose $\mathcal{T}_{ij}^{k+1}$ to be long enough so that the total room for adjustment, which we use $\sum_{s \in \mathcal{T}_{ij}^{k+1}} \lambda_{s,ij}^{\zeta,D}(1 - \lambda_{s,ij}^{\zeta,D})$ as a (conservative) proxy, is on the order of b, a parameter of the heuristic which needs to be carefully chosen (we will call it as the *batch size* parameter). Note that since the optimal deterministic demand can be non-stationary, the number of periods in a batch differs across different batches.

Given the above definition of batches, we now specify how the variability correction quantity in batch k (i.e. $\bar{\delta}_{ij}^k$) is allocated among the different periods in batch $k+1$. For any period t, we compute the new target demand rate as

$$\tilde{\lambda}_{t,ij} = \textsc{Proj}_{[0,1]}\left(\lambda^{\zeta,D}_{t,ij} - \epsilon - u_{t,ij}\cdot \bar{\delta}_{ij}^{\kappa_{ij}(t)-1}\right), \tag{15}$$

where we define

$$u_{t,ij} := \lambda^{\zeta,D}_{t,ij}\left(1-\lambda^{\zeta,D}_{t,ij}\right)\cdot\left[\sum_{v\in\mathcal{T}_{ij}^{\kappa_{ij}(t)}} \lambda^{\zeta,D}_{v,ij}(1-\lambda^{\zeta,D}_{v,ij})\right]^{-1} \tag{16}$$

The above adjustment scheme makes sure that the perturbation in each period is proportional to the *maximum room of adjustment* in the corresponding period: When $\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}\geq 0$ (resp. $\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}<0$), the maximum room of adjustment in period t is $\lambda^{\zeta,D}_{t,ij}-0=\lambda^{\zeta,D}_{t,ij}$ (resp. $1-\lambda^{\zeta,D}_{s,ij}$) since we plan to decrease (resp. increase) the demand rate in batch $\kappa_{ij}(t)$. If the projection operator is inactive for all $s\in\mathcal{T}_{ij}^{\kappa_{ij}(t)}$, it is straightforward to check that the cumulative adjustment in batch $\kappa_{ij}(t)$ is exactly $-\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}$ (since the sum of $u_{t,ij}$ over t within the same batch equals one). Such design ensures that the demand rate adjustment in a period does not exceed the maximum room of adjustment in that period, and is important for the analysis of our dynamic pricing control, *Arc Balancing Control* (ABC), which we formally define next.

Arc Balancing Control

Input: Tuning parameters ϵ,ζ,b

Step 1: Solve **DCP(ζ)** to obtain $\{\lambda^{\zeta,D}_{t,ij}\}_{t,ij}$. Compute batches according to (14).

Step 2: For $t=1$ to T, do:

a. For each i and j, if $t=\chi_{ij}(\kappa_{ij}(t))$:
   - For all $s\in\mathcal{T}_{ij}^{\kappa_{ij}}$, compute $\tilde{\lambda}_{s,ij}$ according to (15)

b. For each i, do:
   - If $C_{t,i}>N$, set $p^{ABC}_{t,ij}=\textsc{Proj}_{t,ij}(p_{t,ij}(\tilde{\lambda}_{t,ij}))$ for all j;
   - Otherwise, set $p^{ABC}_{t,ij}=\bar{p}_{t,ij}$.

The performance of ABC critically depends on the batch size parameter b. This parameter essentially controls the size of each batch. Moreover, since the values of the baseline demand rates $\lambda^{\zeta,D}_{v,ij}$ are all between 0 and 1, the value of b also determines the amount of cumulative stochastic demand variability within a batch. In particular, b needs to be tuned together with ϵ since ϵ governs the size of the buffer. If b is too large, then the cumulative randomness in each batch is too big and we still need to choose a large buffer ϵ to mitigate demand blockage; in this case, the variability correction effect is not significant enough to reduce the loss of static control. On the other hand, if b is too small, then the perturbation in each period will be (with high probability) too big such that the target demand rate in each period cannot be achieved; in other words, the projection operator in (15) becomes active. It turns out that, under proper parameter tuning, the proposed adjustment scheme leads to an asymptotic performance that improves the static pricing controls. The following result characterizes the performance of ABC.

Theorem 3. *Set* $b=(\underline{\tau}^n)^{2/3}$ *and* $\zeta/2=\epsilon=(\underline{\tau}^n)^{-1}(1+16\sqrt{2(b+1)\log T^n})$. *There exists some constant* $\tilde{\Psi}_2$ *independent of* n *such that*

$$\mathcal{L}^{ABC}(n)\leq\tilde{\Psi}_2\cdot\frac{\log(1+n)^{3/2}}{n^{2/3}}.$$

Our result shows that ABC has a significantly better asymptotic performance compared to SPC. As we have discussed above, the improved theoretical performance relies on the fact that the randomness in resource consumption is controlled by our price adjustment scheme, which allows for a smaller buffer size to ensure the same level of demand blockage. Similar to SPC, ABC reserves at least $\epsilon\cdot\underline{\tau}^n$ servers unused in every node to absorb the potential demand variability, and this variability buffer only needs to absorb total demand variability across $\Theta(b)$ periods. Therefore, it suffices to set $\epsilon\approx\sqrt{b}/\underline{\tau}^n=\tilde{\Theta}(n^{-2/3})$, which is a considerable reduction in the buffer size compared to $\epsilon=\tilde{\Theta}(n^{-1/2})$ in Theorem 2; this allows us to reduce the revenue loss caused by the biased static base price due to the need for buffering.

One important implication of Theorem 3 is that it confirms the existence of a policy whose revenue is $O(n^{-2/3})$ away from the deterministic benchmark. This implies that the deterministic benchmark must also be at most $O(n^{-2/3})$ away from the optimal policy which, together with the lower bound result in Theorem 1, further implies that the revenue gap between the optimal policy and any static policy is at least $\Omega(n^{-1/2})$. Thus, we have established that, compared to static policies, our adaptive policy leads to a much smaller revenue gap relative to the *optimal policy*.

Remark 3. The $n^{-2/3}$ relative loss bound for dynamic pricing controls has also appeared in other related settings, i.e., Kim et al. (2018) and Varma et al. (2022). In the settings of these two papers, it has further been shown that $n^{-2/3}$ is the best performance that can be attained in a wide class of dynamic pricing controls. There are two key differences between our paper and these two papers. First, we focus on a transient control problem whereas these two papers assume the system is in steady state and analyze the steady-state control problem. Second, our model is a loss system where demand is lost if there is no server available to serve demand. In contrast, the models in Kim et al. (2018) and Varma et al. (2022) are delayed systems where demand can "wait" in a queue if not matched directly with an available resource and incur some form of waiting cost. Therefore, our results do not follow from their models and analyses. It is also worth noting that the technique used to show that $n^{-2/3}$ is the best achievable performance relies heavily on the steady-state behavior of the delay; as a result, it is not clear to us whether $n^{-2/3}$ is also the best achievable performance in our setting.

Remark 4. In ABC, due to the adaptive correction of stochastic demand variability, the order of cumulative demand variability no longer follows the conventional square-root scaling implied by the Central Limit Theorem. As a result, it is sufficient to keep a smaller buffer size in the order of $n^{-2/3}$ to ensure a small blocking probability compared to the $n^{-1/2}$ in SPC. This deviation from the conventional square root staffing law in queueing theory has also appeared in Besbes et al. (2021a), but for very different reasons. Besbes et al. (2021a) has shown that for ride-hailing systems, the firm should have an extra $n^{2/3}$ capacity than demand to ensure a good quality of service compared to the conventional $n^{1/2}$ staffing rule in the queueing literature; in their context, such deviation is driven by the negative correlation between driver density and the time to reach the riders. Thus, our work complement their work by showing a different mechanism which results in a deviation from the conventional square root staffing rule.

Remark 5. There is a recent line of work which looks into dynamic resource allocation problems and proposes heuristic controls with much tighter loss bound (in the order of n^{-1} in terms of the performance metric in our paper), e.g., Arlotto and Gurvich (2019), Bumpensanti and Wang (2020), Vera et al. (2020) and (2021), and Wang and Wang (2022). It is worth pointing out that while the setting they consider also involves resource constraints, the nature of their resource constraints are very different from ours. In these papers (except for Vera et al. (2020)), there is a fixed amount of resources and each unit of resource can be used no more than *once* throughout the planning horizon. As a result, the resource constraints are effectively only on the *cumulative consumption* of the resources at the end of the planning horizon. In contrast, in our paper, each unit of resources could potentially be used multiple times throughout the planning horizon, so the resource availability is not monotonically decreasing over time and the resource constraints are imposed on the whole sample path rather than just towards the end of the planning horizon. Vera et al. (2020) studies a generalization of NRM model where, besides the initial resource units, additional resource units can arrive over time, so the resource availability is also not monotonically decreasing over time. However, their paper maintains the assumptions that each resource unit can only be used once throughout the planning horizon, and the arrival of additional resource units follows an *exogenous* process. In contrast, the arrival process of resource units in our context is endogenous and resource units are reusable: The pricing decisions affect when and where resource units get utilized which in turn determines when and where these resource units will become available to use again. In sum, the reusable nature of the resources in our model is one of the features which differentiate ride-hailing systems from the conventional network revenue management applications such as airline pricing, and it introduces nuances in the analysis. It is not clear to us whether it is possible (if so, how) to develop a pricing heuristic with sub-$n^{-2/3}$ loss in our setting, and we leave this as a future research direction.

5.2. Outline of the proof of Theorem 3

Let $\pi=$ABC. Define $\eta=(\epsilon\underline{\tau}^n-1)/4$. Similar to the proof of Theorem 2, we divide the analysis into two different cases depending on whether the scaling parameter n is large enough.

Analysis of small n

By definition, we know that as $n\to\infty$, $b=(\underline{\tau}^n)^{2/3}\to\infty$ and

$$\zeta=\frac{2}{\underline{\tau}^n}+\frac{32\sqrt{2(b+1)\log T^n}}{\underline{\tau}^n}\leq 64\cdot\left(\frac{1}{\underline{\tau}^n}+\frac{\sqrt{\log T^n}}{(\underline{\tau}^n)^{2/3}}\right)\leq 128\cdot\frac{\sqrt{\log T^n}}{(\underline{\tau}^n)^{2/3}}\to 0 \tag{17}$$

$$\frac{2\eta}{b}=\frac{8\sqrt{2(b+1)\log T^n}}{b}\leq 16\sqrt{\frac{\log T^n}{b}}=16\sqrt{\frac{\log T^n}{(\underline{\tau}^n)^{2/3}}}\to 0 \tag{18}$$

Define $\Omega:=\max\{n\in\mathbb{Z}_{++}:(\underline{\tau}^n)^{2/3}\leq\max\{4N^2/T^n,256\log T^n,128\sqrt{\log T^n}/\bar{\zeta}\}\}$ (if the right hand side is an empty set, let $\Omega:=0$). Then, similar to the proof of of Theorem 2, for all $n\leq\Omega$,

$$\mathcal{L}^\pi(n)\leq N^2 r_{\max}\leq M_1\frac{(\log T^n)^{2/3}}{(\underline{\tau}^n)^{2/3}}$$

where the constant $M_1:=N^2 r_{\max}\max\{4N^2,256,128/\bar{\zeta}\}$ is independent of n, T, and τ.

Analysis of large n

Our proof for the case with large n proceeds with two major steps. We will first show that, conditioning on a carefully defined "good" event, the prices for all arcs throughout the entire horizon can be explicitly characterized. Moreover, the probability that such event happens is controlled by a tuning parameter of ABC. In the second step, we bound the revenue loss by primarily focusing on what happens in the good event.

When $n>\Omega$, by the definition of Ω, (17) and (18), the following condition holds:

$$\mathbf{C2}:\quad \bar{\zeta}>\zeta=2\epsilon,\ \text{and}\ \frac{2\eta}{b}<1.$$

Define the sets $\mathcal{A}_{ijk}$ and $\mathcal{A}$ as follows:

$$\mathcal{A}_{ijk}:=\left\{\max_{t\in\mathcal{T}_{ij}^k}\left|\sum_{s=\min\{v\in\mathcal{T}_{ij}^k\}}^{t}\delta_{s,ij}^n\right|<\eta\right\}$$

and $\mathcal{A}:=\cap_{i,j,k}\mathcal{A}_{ijk}$. It can be shown that, on event $\mathcal{A}$, condition **C2** ensures that the supply at all of the nodes will not run out throughout the horizon, and the actual adjusted prices are set such that we are able to completely achieve the target demand rate adjustment quantity in each period. The probability of event $\mathcal{A}$ is also directly controlled by the batch size parameter b. We state a lemma that formalizes the above discussions.

LEMMA 4. *If **C2** holds, then for all sample paths on $\mathcal{A}$, the following condition holds for all t,*

$$\mathbb{H}_t: \; C^n_{t,i} \geq N, \forall i, \text{ and } p^\pi_{t,ij} = p^n_{t,ij}(\hat{\lambda}_{t,ij}), \text{ where } \hat{\lambda}_{t,ij} = \lambda^{n,\zeta,D}_{t,ij} - \epsilon - u_{t,ij} \cdot \bar{\delta}^{\kappa_{ij}(t)-1}_{ij} \in (0,1), \forall i,j.$$

Moreover, $\mathbf{P}(\mathcal{A}) \geq 1 - 2N^2(bT^n)^{-1}$.

Lemma 4 is the analogue of the Lemma 3 in the analysis of SPC. Since price of each arc is computed independently, we can analyze the evolution of server utilization on each arc independently (conditioning on $\mathcal{A}$). However, the analysis here is more involved since the prices are adjusted adaptively over time. In particular, we show that the maximum amount of random variability in the utilization of servers on each arc has the same order as the total demand variability within one batch. Given our definition of batches, it can be shown that the total demand variability within a batch is on the order of b. Since b is chosen to be on a smaller order of n, it is on a significantly smaller order compared to T^n.

We now proceed to bound the revenue loss. Since the target demand rate equals $\hat{\lambda}_{t,ij}$ defined in Lemma 4 on event $\mathcal{A}$, following a similar argument as in (12)

$$\mathcal{J}(n,\zeta) - \mathbf{E}^\pi\left[R^\pi(n)\right] \leq \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\{\mathbf{E}^\pi\left[r^n_{t,ij}(\lambda^{n,\zeta,D}_{t,ij}) - r^n_{t,ij}(\hat{\lambda}_{t,ij})\middle|\mathcal{A}\right] + r_{\max}\mathbf{P}(\mathcal{A}^c)\right\} \tag{19}$$

Applying Taylor's expansion to the RHS of (19) and using assumption A3 yields

$$\begin{aligned}
&\sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[r^n_{t,ij}(\lambda^{n,\zeta,D}_{t,ij}) - r^n_{t,ij}(\hat{\lambda}_{t,ij})\middle|\mathcal{A}\right] \\
&= \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[r^n_{t,ij}(\lambda^{n,\zeta,D}_{t,ij}) - r^n_{t,ij}\left(\lambda^{n,\zeta,D}_{t,ij} - \epsilon - u_{t,ij}\cdot\bar{\delta}^{\kappa_{ij}(t)-1}_{ij}\right)\middle|\mathcal{A}\right] \\
&\leq \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[(r^n_{t,ij})'(\lambda^{n,\zeta,D}_{t,ij})\cdot\left(\epsilon + u_{t,ij}\cdot\bar{\delta}^{\kappa_{ij}(t)-1}_{ij}\right) + \Psi_3\cdot\left(\epsilon^2 + \left(u_{t,ij}\cdot\bar{\delta}^{\kappa_{ij}(t)-1}_{ij}\right)^2\right)\middle|\mathcal{A}\right] \\
&\leq \Psi_3\cdot T^n N^2(\epsilon+\epsilon^2) + \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}(r^n_{t,ij})'(\lambda^{n,\zeta,D}_{t,ij})\cdot u_{t,ij}\cdot\mathbf{E}^\pi\left[\bar{\delta}^{\kappa_{ij}(t)-1}_{ij}\middle|\mathcal{A}\right] \\
&\qquad + \Psi_3\cdot\sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[\left(u_{t,ij}\cdot\bar{\delta}^{\kappa_{ij}(t)-1}_{ij}\right)^2\middle|\mathcal{A}\right]
\end{aligned} \tag{20}$$

Since $\delta_{t,ij}$'s are independent random variables with zero mean, we have $\mathbf{E}^\pi[\bar{\delta}^k_{ij}] = 0$. Moreover, since $\lambda^{n,\zeta,D}_{t,ij}\cdot(1-\lambda^{n,\zeta,D}_{t,ij}) \geq \zeta/2$ and $\delta_{t,ij} \leq 1$ almost surely, we know that $|\bar{\delta}^k_{ij}| \leq |\mathcal{T}^k_{ij}| \leq 2(b+1)/\zeta$. Therefore, the second term on the RHS of (20) can be bounded as follows:

$$\begin{aligned}
\left|u_{t,ij}\cdot\mathbf{E}^\pi\left[\bar{\delta}^k_{ij}\middle|\mathcal{A}\right]\right| &= \left|u_{t,ij}\mathbf{P}(\mathcal{A})^{-1}\left\{\mathbf{E}^\pi\left[\bar{\delta}^k_{ij}\right] - \mathbf{E}^\pi\left[\bar{\delta}^k_{ij}\middle|\mathcal{A}^c\right]\mathbf{P}(\mathcal{A}^c)\right\}\right| \\
&= \left|u_{t,ij}\mathbf{P}(\mathcal{A})^{-1}\mathbf{E}^\pi\left[\bar{\delta}^k_{ij}\middle|\mathcal{A}^c\right]\mathbf{P}(\mathcal{A}^c)\right| \leq \frac{4N^2}{b\zeta T^n(1-2N^2(bT^n)^{-1})}
\end{aligned} \tag{21}$$

where the first equality holds by law of total expectation, and the inequality follows by Lemma 4, $u_{t,ij} \le 1/(b+1)$ and $\mathbf{E}^{\pi}[\bar{\delta}_{ij}^{k}|\mathcal{A}^c] \le 2(b+1)/\zeta$. For the third terms on the RHS of (20), we have:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^{\pi}\left[\sum_{t=1}^{T^n}\left(u_{t,ij}\cdot\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}\right)^2\middle|\mathcal{A}\right] \le \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{1}{b}\right)^2\cdot\mathbf{E}^{\pi}\left[\sum_{t=1}^{T^n}\left(\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}\right)^2\middle|\mathcal{A}\right]$$
$$\le \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T^n}\frac{1}{b^2}\cdot\eta^2 = \frac{T^n N^2}{b^2}\cdot 32(b+1)\log T^n \le \frac{64 T^n N^2\log T^n}{b} \tag{22}$$

where the first inequality holds by the definition of $u_{t,ij}$ and $\mathcal{T}_{ij}^{k}$, the second inequality holds by **C2**, and the equality holds by the definition of η.

Combining (19), (20), (21), (22) and Lemma 4 together, we have

$$\begin{aligned}
\mathcal{L}^{\pi}(n) &= (\mathcal{J}(n,0)-\mathcal{J}(n,\zeta)+\mathcal{J}(n,\zeta)-\mathbf{E}^{\pi}[R^{\pi}(n)])/T^n \\
&\le \Psi_0\zeta + N^2\left[\Psi_3\left(\epsilon+2\epsilon^2+\frac{4N^2}{b\zeta T^n(1-2N^2(bT^n)^{-1})}+\frac{64\log T^n}{b}\right)+\frac{2N^2 r_{\max}}{b}\right] \\
&\le \left[2\Psi_0+3\Psi_3N^2\right]\cdot\epsilon+\frac{2N^2}{b}\left[\Psi_3\left(\frac{N^2}{1-2N^2(bT^n)^{-1}}+32\log T^n\right)+N^2 r_{\max}\right] \\
&\le \frac{2\Psi_0+3\Psi_3N^2}{\underline{T}^n}+\left[32\Psi_0+48N^2\Psi_3+2N^2\Psi_3(2N^2+32\log T^n)+2r_{\max}N^4\right]\cdot\frac{\sqrt{\log T^n}}{(\underline{T}^n)^{2/3}} \\
&\le M_2\cdot\frac{1+(\log T^n)^{\frac{3}{2}}}{(\underline{T}^n)^{2/3}}
\end{aligned}$$

where $M_2 = 32\Psi_0+112N^2\Psi_3+4N^4\Psi_3+2N^4 r_{\max}$ is independent of n, the first inequality follows by Lemma 2, the second inequality holds since $\zeta T^n \ge 2$, the third inequality follows by (18) and the definition of Ω, and the last inequality follows since $\underline{T}^n > 1$. Setting $\tilde{\Psi}_2 = \max\{M_1, M_2\}[(\log(2))^{-3/2} + (1+\log(T)/\log(2))^{3/2}]\underline{T}^{-2/3}$ completes the proof. □

6. Numerical Studies

To test the empirical performance of our heuristic controls, we now conduct two sets of numerical studies. In the first set, we use synthetic data to examine the performance of different heuristic controls under different model parameters. In the second set, we calibrate our model using a real-world taxi ride data set to study the practical impact of dynamic pricing. Additional details of the numerical studies can be found in Appendix EC.3.

6.1. Synthetic Data

We generate a synthetic network with 5 nodes. The travel time between two nodes depends on the "distance" as well as the travel direction: We first generate an integral Euclidean distance matrix, and then set the actual travel time to be the distance multiplied by a random congestion factor (uniformly distributed between $[0.5, 1.5]$). In the unscaled problem, the travel times range from 2 to 10 decision periods, and the length of planning horizon is $T = 30$ decision periods. We assume a linear demand function. In particular, we set $\lambda_{t,ij}(p_{t,ij}) = \mu_{ij}(\alpha - \beta_{ij} \cdot p_{t,ij})$, where μ_{ij} can be interpreted as the "market size" of each route, the price sensitivity is calculated as $\beta_{ij} = \beta/\tau_{ij}$ to capture the fact that travelers are willing to pay more for longer trips. The hyper-parameters $\boldsymbol{\mu}, \alpha$, and β are generated randomly. The initial resource distribution is set to be roughly balanced based on Little's Law, i.e., we set $C_i \approx \gamma_i \cdot \sum_{j=1}^{N} \mu_{ij} \alpha \tau_{ij}$, where γ_i can be interpreted as a load factor which governs the scarcity of supply.

We evaluate the performance of SPC and ABC using Monte Carlo simulations. We observe that, in general, tuning the value of ζ will not improve the performance of the controls by much. Therefore, in all of the experiments, we only report the performance under $\zeta = 0$. Our choices of ϵ and b are determined using grid search. In particular, we set ϵ and b to be some tuning constants multiplied by the order of n defined in Theorems 2 and 3. Note that in our numerical studies, we do not directly use the exact expressions of parameters defined in Theorems 2 and 3. While these expressions are sufficient to establish our theoretical results, they may not be optimal. Thus, for practical implementation, these parameters should be chosen using simulation to get the best empirical performance. Additional details of the numerical study and results in this subsection can be found in Appendix EC.3.1.

6.1.1. Performance under varying n. We first simulate the performance of our heuristic controls for varying n: Recall that a higher n corresponds to an urban area with heavier traffic. For expositional simplicity, we only report the results for one set of simulation parameters (i.e., $\boldsymbol{\mu}, \alpha$, β, and $\boldsymbol{\gamma}$) since the findings are robust to these parameters. Figure 1 plots the losses (multiplied by $\sqrt{n}$) for our proposed pricing controls for various n. The non-increasing trends for both curves in Figure 1 confirm that both SPC and ABC are asymptotically optimal, and their asymptotic order of losses are no bigger than $n^{-1/2}$. In particular, the loss of SPC is on the order of $n^{-1/2}$, whereas the loss of ABC is noticeably smaller than $n^{-1/2}$. In terms of total expected revenues, the improvement of ABC over SPC ranges from 1.88% to 2.66%. The detailed performance of all heuristic controls can be found in Table 1. (Columns "Loss %" report the percentage loss compared with $\mathcal{J}(0, n)$; columns "Loss \$" report the per period loss as defined in (9); columns "Rev. \$" report the expected revenue; and, column "Rev. Incr. %" reports the percentage revenue improvement of ABC over SPC.) These results highlight the benefit of dynamic pricing for managing supply and demand in ride-hailing networks.

Upon a closer inspection of the simulation results, we find that dynamic pricing facilitates a more balanced distribution between supply and demand, which helps increase the number of total

n	SPC Loss \$	SPC Loss %	SPC Rev \$	SPC Admit #	ABC Loss \$	ABC Loss %	ABC Rev. \$	ABC Rev. Incr. %	ABC Admit Incr. %	ABC Avg. Price Incr. %
40	8.76	7.40	131523	772	5.85	4.94	135020	2.66	1.87	0.77
80	6.50	5.49	268478	1577	3.66	3.09	275299	2.54	1.70	0.82
120	5.34	4.52	406872	2390	2.60	2.19	416760	2.43	1.76	0.66
160	4.66	3.94	545762	3205	2.10	1.77	558073	2.26	1.67	0.57
200	4.10	3.47	685564	4026	1.68	1.42	700108	2.12	1.30	0.81
240	3.77	3.18	825113	4846	1.42	1.20	841966	2.04	1.36	0.67
280	3.33	2.81	966307	5675	1.22	1.03	984035	1.83	1.18	0.64
320	3.21	2.71	1105496	6492	1.05	0.89	1126183	1.87	1.01	0.85
360	3.13	2.64	1244541	7309	0.89	0.75	1268690	1.94	1.14	0.79
400	2.94	2.48	1385085	8134	0.77	0.65	1411184	1.88	1.19	0.68

Table 1 Performances of proposed pricing controls under varying n

admitted demands in spite of an increase in average price; see Table 1 again. (Column "Admit #" report the total number of admitted customers; columns "Admit Incr. %" report the percentage increment in the total number of admitted customers under ABC over that under SPC; column "Avg Price Incr. %" report the percentage increment in the average price charged to an admitted customer under ABC over that charged under SPC). More interestingly, it can be observed that, under ABC, the increase in the average price is noticeably smaller than the revenue improvement over SPC, which suggests that the revenue improvement under ABC comes mainly from admitting more customers instead of from charging higher prices. This has an immediate practical implication: Despite our original focus on improving revenue, dynamic pricing also helps in achieving one important goal of many ride-hailing systems, i.e., to increase the number of customers served.

In addition to the linear demand models, we also test the performances of our proposed controls under exponential and logit demand models; the findings are qualitatively similar to the case of linear demand model. We also tested the robustness of the proposed controls with respect to the value of T. The results are summarized in Section EC.3.1.

6.1.2. Benefit of the intertemporal feature of pricing. We would like to point out an important observation: While the demand parameters in our synthetic data set are set to be stationary over time under the choices of parameters above, the optimal deterministic demand $\lambda_{t,ij}^{\zeta,D}$ is *not* necessarily stationary over time for any feasible choice of ζ due to the beginning and end-of-horizon effects, which are unavoidable in practise since urban traffic is cyclic in nature. In fact, the non-stationarity in optimal deterministic demand rate is significant even if we assume a fairly long horizon during which the demand function is stationary: If we set $T = 300$ (i.e., 30 times larger than the maximum travel time), we observe that the optimal deterministic demand is still not stationary for a significant portion of time (see Figure 2 for the optimal deterministic demand

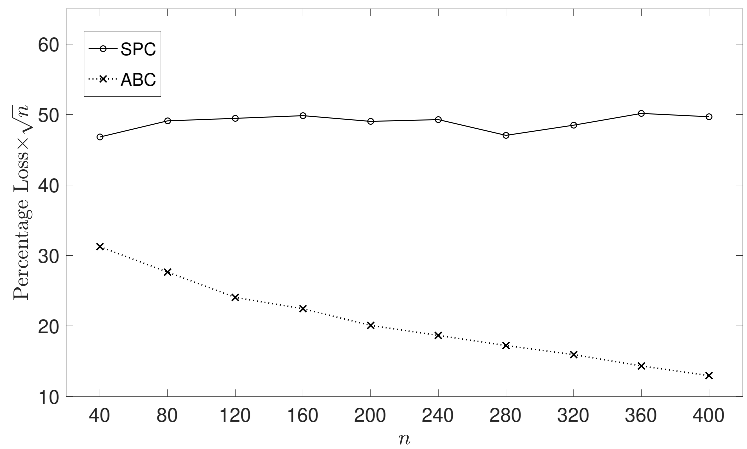
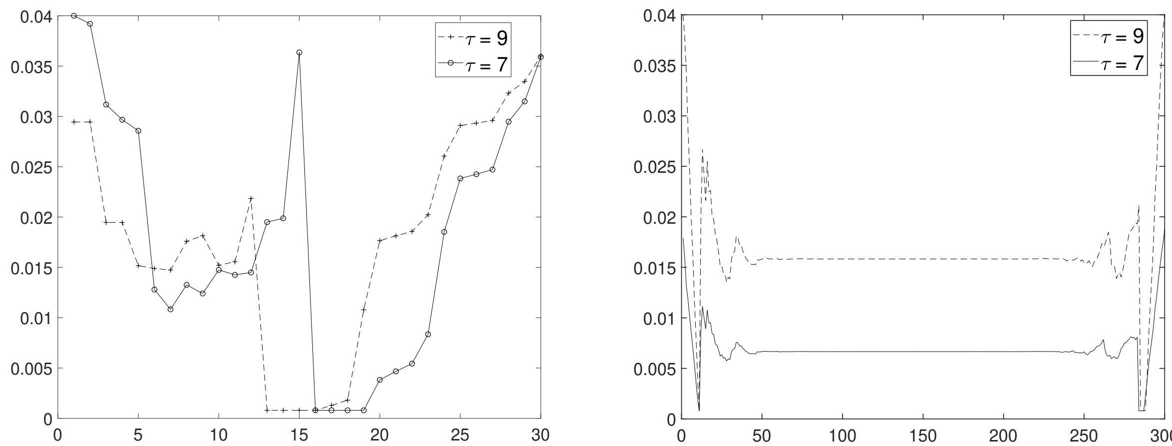


Figure 1 $\sqrt{n}$- scaled losses under varying n

trajectory on two O-D pairs for both $T = 30$ and $T = 300$). To put the numbers in perspective, under the choice of the experiment, even if the maximum travel time is as short as 10 minutes, setting $T = 300$ requires demand function to be stationary over a 5-hour window, which is unlikely to hold in reality. This observation implies that dynamic pricing controls that only use stationary price as baseline (as opposite to non-stationary but static baseline) could have a poor performance. In fact, when we simulate the performance of our proposed pricing controls under a stationary baseline price, i.e., restricting the optimal deterministic solution to be stationary over time, we found that as n increases, both of the modified heuristics are *not* asymptotically optimal anymore (see Table EC.1 in the Appendix for more details). In fact, our numerical results suggest that the stationarity restriction on baseline prices renders an additional revenue loss of 3% to 7% for both SPC and ABC when compared with the heuristics that use optimal non-stationary (but static) baseline price. This observation highlights the importance of the intertemporal feature of pricing in non-stationary settings, especially for urban transport where travel patterns tend to fluctuate throughout the day.

6.2. Manhattan Yellow Taxi Data

To quantify the performance of our heuristic controls in a more realistic setting, we calibrate a model using yellow cab data obtained from NYC Taxi & Limousine Commission (New York City 2020). We use the Manhattan's yellow taxi trip data from January to December in 2019 and limit our attentions to weekdays from 7 a.m. to 4 p.m. (see similar choice in Buchholz (2022)). After cleaning the incomplete data entries, we obtain a data set of roughly 20 million records. In the original data set, the origin and destination of each trip is specified by the index of TLC Taxi Zone, defined in New York City (2020). For tractability issue, we further combine adjacent zones



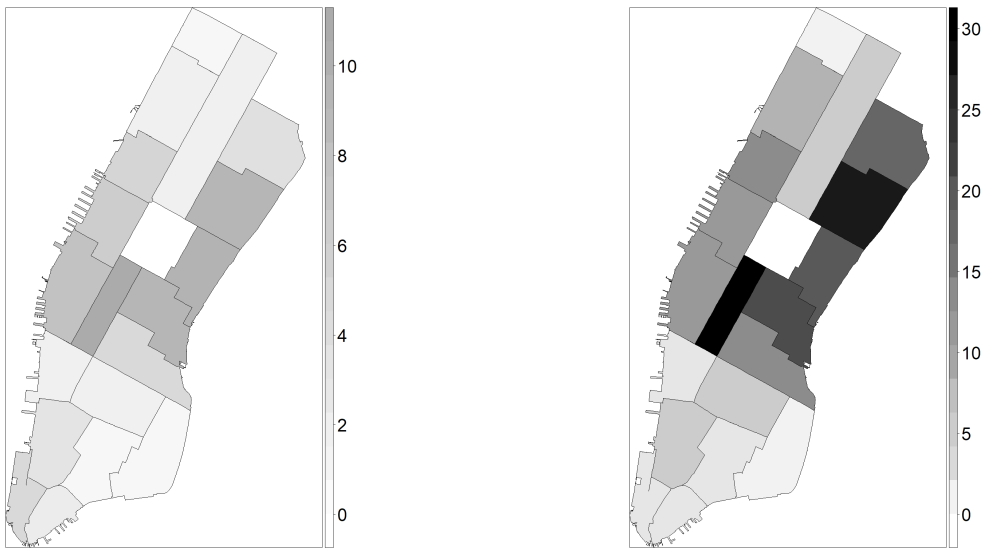
(a) $T = 30$ (b) $T = 300$

Figure 2 Optimal deterministic demand path under different T

to make the total number of regions to 20 (see Table EC.5 in Appendix for the detailed zoning method). We estimate the potential market size and travel time for each O-D pair for each 10-min interval from the data set. Similar to the synthetic experiment, we use a linear demand function where $\lambda_{t,ij}(p_{t,ij}) = \mu_{ij}(\alpha - \beta_{ij} \cdot p_{t,ij})$ to fit the data. Specifically, μ_{ij} represents the market size, and we compute it by multiplying the average number of trips from region i to region j by a constant that is greater than 1. The median of the travel time ranges from 258 seconds to 2,282 seconds, whereas the average number of trip per 10-minute interval ranges from 0 to 59.6.

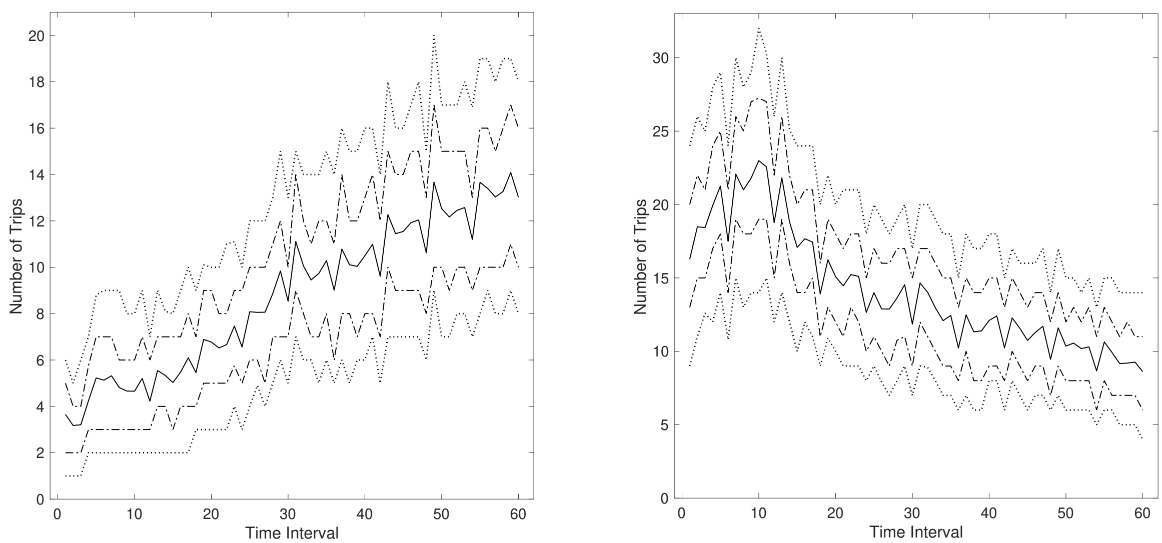
We first present some traffic patterns revealed from the data. Figure 3 illustrates the heat-maps of the average number of trips from region 13 (midtown center region) for two different 10-minute intervals using the same color coding. (Region 13 is represented as the white square in the center of the map.) As the graph clearly illustrates, the number of trips is heterogeneous along both geographical and time dimensions. Figure 4 provides summary statistics for the number of trips between region 13 and region 15 for each 10-minute interval, where we use the solid line represents the average value, the dashed line to represent the 25% and 75% percentile, and dotted line to represent the 10% and 90% percentile. Within the same day, the number of trips from Lincoln Square region to midtown center region decreases whereas the number of trips from midtown center region to Lincoln Square region increases. This is consistent with the workday commute patterns.

For the simulation, we set the planning horizon to be 10 hours (i.e., from 7 a.m. to 4 p.m) and set the length of one period to be 5 seconds (this guarantees that no more than one customer arrive within the same period for each O-D pair). As a result, $T = 7,200$. Given that the market size is estimated for every 10-minute interval, the demand function is assumed to be piece-wise stationary during each 10-minute interval. For tractability issue, we assume that the optimal baseline static



(a) 8:00 AM - 8:10 AM (b) 3:00 PM - 3:10 PM

Figure 3 Number of trips from region 13



(a) From 13 to 15 (b) From 15 to 13

Figure 4 Number of Trips for each 10-minute Interval

price is also piece-wise stationary. In particular, when solving **DCP**, we require that the optimal demand rate to be stationary for each 5-minute interval (see Appendix EC.3.2 for a formal definition). The initial distribution of taxis $\{C_i\}_{i=1}^N$ is computed similarly with the synthetic experiment. We simulate all the proposed controls under different scenarios: The "high-congestion" scenario, where we choose the 75-percentile travel time and number of pickups; the "average-congestion" scenario, where we choose the median travel time and the average number of pickups; and the “low-congestion” scenario, where we choose the 25-percentile travel time and number of pickups. Table 2 summarizes the performances of all heuristic controls.

As predicted by our theoretical analysis, the performance improves as the flexibility in pricing control increases. Specifically, ABC reduces the loss by more than a half compared to SPC in all three congestion scenarios, and offers a revenue improvement which ranges from 5.0% to 6.4%. This suggests that the benefit of dynamic pricing can be quite significant. We also observe that the increment in revenue does not sacrifice the service level. In fact, ABC results in an increase of admitted demand between 3.0% to 4.1% when compared to SPC.

Traffic	SPC			ABC		
Condition	Loss %	Admit	Avg Price	Loss %	Admit	Avg Price
High	8.33	63404	11.20	3.74	65288	11.42
Average	9.66	50362	10.62	4.15	52441	10.82
Low	10.70	35101	9.84	4.93	36430	10.09

Table 2 Performances of proposed pricing controls under different traffic conditions

7. Extension: Joint Control of Pricing and Server Relocation

In this section, we extend our model to a setting where the firm can jointly control the prices of rides and the relocation of servers that are not currently in use (for convenience, we will call them *empty* servers from now on). Similar to existing literature, we assume that there is a marginal cost of relocating one unit of supply from node i to j in period t, denoted by $c_{t,ij}$ (see e.g., Hosseini et al. 2021, Benjaafar et al. 2021, and Akturk et al. 2021). We assume that the time needed to relocate a server from node i to j is the same as the travel time from i to j, and that an empty server cannot take any customer during the process of relocating to a different node. Below, we first give a formal definition of the model, and then discuss how to generalize ABC to handle the relocation decision.

7.1. Model Formulation

Let $Q_{t,ij}$ denote the number of servers being relocated along arc $i \to j$ in period t. In period t, the following sequence of events take place. The firm first reviews the number of empty servers at each node i, and then decides the prices $p_{t,ij}$ for all arcs. After observing the arriving demand in period t, the firm further decides how many empty servers to relocate from node i to node j for all $i \neq j$. (From the perspective of performance analysis, it does not matter whether these two decisions are made sequentially or simultaneously.) The firm's goal is to find an admissible *joint pricing and* *relocation control* to maximize its cumulative expected profit during the horizon, defined as the total revenues collected from serving customers minus the total relocation costs.

Let Π denote the class of admissible controls. Then, the firm's spatial dynamic optimization problem can be formulated as follows:

$$(\textbf{R-SDP}) \quad \mathcal{J}^{R,\star} := \max_{\pi\in\Pi} \quad \mathbf{E}^{\pi}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} p^{\pi}_{t,ij} D_{t,ij}(p^{\pi}_{t,ij}) - c_{t,ij} Q^{\pi}_{t,ij}\right] \tag{23}$$

$$\text{s.t.} \quad C_i - \sum_{j=1}^{N}\sum_{s=1}^{t}\left(D_{s,ij}(p^{\pi}_{s,ij}) + Q^{\pi}_{s,ij}\right) + \sum_{j=1}^{N}\sum_{s=1}^{t-\tau_{ji}}\left(D_{s,ji}(p^{\pi}_{s,ji}) + Q^{\pi}_{s,ji}\right) \geq 0, \ \forall i,t \tag{24}$$

$$p^{\pi}_{t,ij} \in \mathcal{P}_{t,ij}, Q^{\pi}_{t,ij} \in \mathbb{N} \ \forall i,j,t \tag{25}$$

In **R-SDP**, the constraints must hold almost surely and the expectation in (23) is taken with respect to the stochastic processes induced by the control π. In comparison to constraint (2), the flow-balance in (24) also need to capture the impact of relocation decisions.

Similar to our base model, we introduce a class of parameterized *deterministic convex optimizations*. For any $\zeta, \zeta^q \in \mathbb{R}_+$, we define

$$(\textbf{R-DCP}(\zeta,\zeta^q)) \quad \mathcal{J}^{R}(\zeta,\zeta^q) := \max_{\{\lambda_t, q_t\}} \quad \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} [r_{t,ij}(\lambda_{t,ij}) - c_{t,ij} q_{t,ij}] \tag{26}$$

$$\text{s.t.} \quad C_i - \sum_{j=1}^{N}\sum_{s=1}^{t}(\lambda_{s,ij} + q_{s,ij}) + \sum_{j=1}^{N}\sum_{s=1}^{t-\tau_{ji}}(\lambda_{s,ji} + q_{s,ji}) \geq 0, \forall i,t \tag{27}$$

$$\zeta \leq \lambda_{t,ij} \leq 1-\zeta, \ q_{t,ij} \geq \zeta^q \ \forall t,i,j \tag{28}$$

where, by convention, $\mathcal{J}^R(\zeta,\zeta^q) = -\infty$ when $\textbf{R-DCP}(\zeta,\zeta^q)$ is not feasible. Note that, in addition to replacing the random demands by their means, we also approximate the *integer*-valued relocation decision $Q_{t,ij}$ by a continuous variable $q_{t,ij}$. By abuse of notation, we will denote the optimal solution to $\textbf{R-DCP}(\zeta,\zeta^q)$ by $\lambda^{\zeta,\zeta^q,D}_{t,ij}$ and $q^{\zeta,\zeta^q,D}_{t,ij}$. Compared with **DCP** in the setting where pricing is the only decision to make, in **R-DCP**, we introduce one more cushion parameter ζ^q for the extra relocation quantity decision. From the technical perspective, similar to the idea discussed in Remark 1, if we do not introduce such parameter (i.e., we simply set $\zeta^q = 0$), then $q^{\zeta,0,D}_{t,ij}$ is likely to equal zero for some t, i, and j (due to the linearity of objective function in relocation quantity), which prohibits us from adaptively adjusting relocation quantity downward when needed.

In the same spirit as the asymptotic scaling defined in Section 2.3, we will investigate the performance of our heuristic control in the setting where we scale up the demand and supply by a common factor n. In particular, while the other scaling parameters are defined in the same way as in (8), we define the relocation cost as $c^n_{t,ij} = c_{\lceil t/n \rceil, ij}$ for all t and $i \to j$. In other words, the relocation cost is also assumed to be piece-wise stationary. Adopting a similar notation, we will call **R-SDP**(n) (resp. **R-DCP**(n, ζ, ζ^q)) the stochastic control problem (resp. deterministic relaxation) for the scaled system with factor n. We further denote by $\mathcal{J}^R(n, \zeta, \zeta^q)$ and $\{\lambda^{n,\zeta,\zeta^q,D}_{t,ij}, q^{n,\zeta,\zeta^q,D}_{t,ij}\}_{t,ij}$ the optimal value and optimal solution of **R-DCP**(n, ζ, ζ^q), respectively.

7.2. Description of the heuristic

We now proceed to introduce our heuristic control, the *Joint Relocation and Arc Balancing Control* (R-ABC). At a high level, R-ABC and ABC share a similar idea of adaptive variability correction: Both of them divide the horizon into small batches and perturb the demand rates such that the cumulative stochastic demand variability in the previous batch is corrected in the current batch. What is different in R-ABC is that both the definition of batches and the adjustment of demand rates (via pricing) and target relocation quantities are affected by the way relocation decisions are made: Since target relocation quantity under the proposed policy may not always be an integer, the relocation decision in R-ABC is made with a natural randomization scheme, which introduces some additional stochastic variability that needs to be carefully accounted for.

In R-ABC, we define two types of batches: the first type for the purpose of adjustment in the target demand rate (we call them the *demand batches*), and the second type for the purpose of adjustment in the target relocation quantity (we call them the *relocation batches*). The demand batches, denoted by $\{\mathcal{T}^k_{ij}\}_{i,j,k}$ are defined in the same way as in ABC. Formally, we continue to use the same definition of $\mathcal{T}^k_{ij}$, $\kappa_{ij}(t)$, $\chi_{ij}(\kappa_{ij}(t))$, and $\bar{\delta}^k_{ij}$ as in Section 5. As for the relocation quantity, we denote by $\delta^q_{t,ij} := Q^\pi_{t,ij} - \mathbf{E}[Q^\pi_{t,ij}]$ the stochastic variability in relocation quantity induced by the randomization scheme which we will define later. We define the relocation batches, denoted by $\{\mathcal{T}^{q,k}_{ij}\}_{i,j,k}$, as a partition of the selling horizon based on $q^{\zeta,\zeta^q,D}_{t,ij}$ and $\delta^q_{t,ij}$. Formally, let $\mathcal{T}^{q,k}_{ij}$ be the k^{th} relocation batch for arc $i \to j$. Define the *cumulative (stochastic) allocation variability* (i.e., the difference between the realized allocation quantity and the targeted relocation quantity) during the k^{th} relocation batch for arc $i \to j$ as $\bar{\delta}^{q,k}_{ij} := \sum_{t \in \mathcal{T}^{q,k}_{ij}} \delta^q_{t,ij}$. We further denote by $\kappa^q_{ij}(t)$ the index of the relocation batch which period t belongs to, and by $\chi^q_{ij}(k) = \arg\max_s \{s \in \mathcal{T}^{q,\kappa_{ij}(t)}_{ij}\}$ the index for the last period in relocation batch k for arc $i \to j$. The relocation batches are defined sequentially as follows: Define $\mathcal{T}^{q,0}_{ij} = \emptyset$ (which implies that $\bar{\delta}^{q,0}_{ij} = 0$) and $\chi^q_{ij}(0) = 0$. At the end of batch batch $k \geq 0$, define the end of next batch $k+1$ as

$$\chi^q_{ij}(k+1) := \min \left\{ \chi^q_{ij}(k) < s \leq T \; : \; \sum_{v=\chi^q_{ij}(k)+1}^{s} q^{\zeta,\zeta^q,D}_{v,ij} \geq b \right\} \tag{29}$$

For any period t, the target relocation quantity is computed adaptively as

$$\tilde{q}_{t,ij} = \left(q_{t,ij}^{\zeta,\zeta^q,D} - \epsilon^q - u_{t,ij}^q \cdot \bar{\delta}_{ij}^{q,\kappa_{ij}^q(t)-1} \right)^+, \text{ where } u_{t,ij}^q := q_{t,ij}^{\zeta,\zeta^q,D} \cdot \left[\sum_{v \in \mathcal{T}_{ij}^{q,\kappa_{ij}^q(t)}} q_{v,ij}^{\zeta,\zeta^q,D} \right]^{-1} \tag{30}$$

where ϵ^q is a buffer parameter that serves a similar purpose to ϵ for demand adjustment. We would like to point out that the relocation batches and demand batches are defined based on similar ideas. The difference between the detailed definitions (in particular, the adjustment coefficients) is driven by the fact that, in contrast to demand rate which is bounded from above by 1, there is no upper limit on the upward adjustment of relocation quantity.

Below, we give the formal definition of the proposed policy.

Joint Relocation and Arc Balancing Control

Input: Tuning parameters $\epsilon, \epsilon^q, \zeta, \zeta^q, b$

Step 1: Solve **R-DCP**(ζ, ζ^q) to obtain $\{\lambda_{t,ij}^{\zeta,\zeta^q,D}, q_{t,ij}^{\zeta,\zeta^q,D}\}_{t,ij}$.

Compute demand and relocation batches using (14) and (29), respectively.

Step 2: For $t = 1$ to T, do:

Price adjustment

a. and b. Same as Step 2.a and 3.b in ABC.

Empty Server Relocation

c. After observing the realized demand, update the inventory level as $\tilde{C}_{t,i} = C_{t,i} - \sum_{j=1}^N D_{t,ij} + \sum_{j=1}^N D_{(t-\tau_{ji})^+,ji}$.

d. For each node $i = 1, \ldots, N$:

For each arc $i \to j$:

- Compute $\tilde{q}_{t,ij}$ according to (30)
- If $\tilde{q}_{t,ij}$ is integer, then compute $\hat{Q}_{t,ij} = \tilde{q}_{t,ij}$; otherwise, sample

$$\hat{Q}_{t,ij} = \begin{cases} \lfloor \tilde{q}_{t,ij} \rfloor & \text{with probability } \lceil \tilde{q}_{t,ij} \rceil - \tilde{q}_{t,ij} \\ \lceil \tilde{q}_{t,ij} \rceil & \text{with probability } \tilde{q}_{t,ij} - \lfloor \tilde{q}_{t,ij} \rfloor \end{cases}$$

- Set $Q_{t,ij}^{\text{R-ABC}} = \min\{\hat{Q}_{t,ij}, \tilde{C}_{t,i}\}$, and then update $\tilde{C}_{t,i} \leftarrow \tilde{C}_{t,i} - Q_{t,ij}^{\text{R-ABC}}$

Update the inventory level as $C_{t+1,i} = \tilde{C}_{t,i}$.

It is straightforward to verify that the sampling distribution defined in Step 2.d. is a valid probability distribution that guarantees that $\mathbf{E}[\hat{Q}_{t,ij}] = \tilde{q}_{t,ij}$. In other words, if there is sufficient number of servers in the origin node, the expected relocation quantity equals to the target relocation quantity. The following theorem characterizes the performance of R-ABC.

Theorem 4. *Set* $b = (\underline{\tau}^n)^{2/3}$ *and* $\zeta/2 = \zeta^q/2 = \epsilon = \epsilon^q = (\underline{\tau}^n)^{-1}(1 + 32\sqrt{(b+1)\log T^n})$. *There exists some constant* $\tilde{\Psi}_3$ *independent of* n *such that*

$$\mathcal{L}^{R-ABC}(n) \leq \tilde{\Psi}_3 \cdot \frac{\log(1+n)^{3/2}}{n^{2/3}}.$$

Theorem 4 confirms that, under a similar choice of control parameters, R-ABC achieves the same asymptotic performance in the joint control setting as ABC does in the pricing-only setting. It also confirms that the batched adjustment scheme can be applied to both demand rate and relocation quantity. Moreover, while Theorem 4 uses a specific set of control parameters, the same asymptotic order can be achieved by requiring $\epsilon + \epsilon^q$ to be on the order of $\Theta(n^{-2/3}\log^{1/2} n)$. (We omit the formal proof since it is straightforward.) In other words, the buffer servers can be secured via *only* the perturbation of the solution of relocation (resp. pricing) decision but not pricing (resp. relocation) decision, i.e., $\epsilon = 0$ and $\epsilon^q = \Theta(n^{-2/3}\log^{1/2} n)$ (resp. $\epsilon^q = 0$ and $\epsilon = \Theta(n^{-2/3}\log^{1/2} n)$).

8. Closing Remarks

In this paper, we have studied a spatial-intertemporal pricing problem in which a firm, which provides ride-hailing service over a network in urban areas, uses pricing as the control to balance supply and demand. Unlike many papers in the literature that focus on steady-state analysis, we focus on transient control and allows demands to be both stochastic and non-stationary. We analyze the performance of static pricing controls and a novel dynamic pricing control. Both our theoretical and numerical results reveal that the benefit of dynamic pricing over static pricing can be significant especially in non-stationary settings. Finally, we have also demonstrated how our price adjustment scheme can be combined with server relocation decisions in a natural way to develop an effective joint relocation and pricing control.

Our work leaves open many interesting future research directions. We briefly discuss some of them. First, we have assumed in this paper that servers who are currently stationed at a given location will travel to another location either because there is a customer request or because the firm decides to relocate the servers. In practice, however, servers may actively and strategically move around from one location to another location in a decentralized fashion. It would be interesting to develop provably near-optimal dynamic pricing controls for such setting, and to also understand the value of centralized relocation control versus a decentralized server relocation. Second, we have currently assumed that price adjustment only affects the size of demand. In practice, the change in price could affect both the size of demand and the number of available servers who are willing to work at the given price. In such case, how should we dynamically adjust the prices over time? Third, while our dynamic policy achieves $O(n^{-2/3})$ revenue loss, it is not clear if it is possible to develop policies with tighter loss bounds, or if the gap between the deterministic upper bound and the optimal policy is already in the order of $n^{-2/3}$ so there is limited room for further improvement of the loss bound; results along either direction is of great theoretical importance. Finally, for tractability, we have assumed that the travel time between two nodes in the network is deterministic. In practice, travel time can be uncertain and cannot always be accurately approximated with a deterministic variable. Since uncertainty in travel times has ripple effects on the distribution of available servers across the network at future times, there is a need to better understand how to do price adjustment in such setting.

References

Afeche, Philipp, Zhe Liu, Costis Maglaras. 2018. Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance. *Columbia Business School Research Paper* .

Akturk, Deniz, Ozan Candogan, Varun Gupta. 2021. Network inventory management: Approximate optimality in large-scale systems. *Available at SSRN 3842817* .

Arlotto, Alessandro, Itai Gurvich. 2019. Uniformly bounded regret in the multisecretary problem. *Stochastic Systems* **9**(3) 231–260.

Ata, Baris, Nasser Barjesteh, Sunil Kumar. 2019. Spatial pricing: An empirical analysis of taxi rides in new york city. Tech. rep., Working Paper.

Bai, Jiaru, Kut C So, Christopher S Tang, Xiqun Chen, Hai Wang. 2019. Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing & Service Operations Management* **21**(3) 556–570.

Balseiro, Santiago R, David B Brown, Chen Chen. 2021. Dynamic pricing of relocating resources in large networks. *Management Science* **67**(7) 4075–4094.

Banerjee, Siddhartha, Daniel Freund, Thodoris Lykouris. 2021. Pricing and optimization in shared vehicle systems: An approximation framework. *Operations Research* .

Banerjee, Siddhartha, Yash Kanoria, Pengyu Qian. 2018. State dependent control of closed queueing networks with application to ride-hailing. *arXiv preprint arXiv:1803.04959* .

Benjaafar, Saif, Zicheng Wang, Xiaotang Yang. 2021. Autonomous vehicles for ride-hailing. *Available at SSRN 3919411* .

Besbes, Omar, Francisco Castro, Ilan Lobel. 2021a. Spatial capacity planning. *Operations Research* .

Besbes, Omar, Francisco Castro, Ilan Lobel. 2021b. Surge pricing and its spatial supply response. *Management Science* **67**(3) 1350–1367.

Besbes, Omar, Adam N Elmachtoub, Yunjie Sun. 2019. Static pricing: Universal guarantees for reusable resources. *Proceedings of the 2019 ACM Conference on Economics and Computation*. 393–394.

Bimpikis, Kostas, Ozan Candogan, Daniela Saban. 2019. Spatial pricing in ride-sharing networks. *Operations Research* **67**(3) 744–769.

Braverman, Anton, Jim G Dai, Xin Liu, Lei Ying. 2019. Empty-car routing in ridesharing systems. *Operations Research* **67**(5) 1437–1452.

Buchholz, Nicholas. 2022. Spatial equilibrium, search frictions, and dynamic efficiency in the taxi industry. *The Review of Economic Studies* **89**(2) 556–591.

Bumpensanti, Pornpawee, He Wang. 2020. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* **66**(7) 2993–3009.

Cachon, Gerard P, Kaitlin M Daniels, Ruben Lobel. 2017. The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management* **19**(3) 368–384.

Castillo, Juan Camilo, Dan Knoepfle, Glen Weyl. 2017. Surge pricing solves the wild goose chase. *Proceedings of the 2017 ACM Conference on Economics and Computation*. 241–242.

Chen, Q, S Jasin, I Duenyas. 2015. Real-time dynamic pricing with minimal and flexible price adjustment. *Manag. Sci.* **62**(8) 2437–2455.

Chen, Y, R Levi, C Shi. 2017. Revenue management of reusable resources with advanced reservations. *Prod. Oper. Manag.* **26**(5) 836–859.

Chen, Yiwei, Ming Hu. 2019. Pricing and matching with forward-looking buyers and sellers. *Manufacturing & Service Operations Management* .

Fang, Zhixuan, Longbo Huang, Adam Wierman. 2019. Prices and subsidies in the sharing economy. *Performance Evaluation* **136** 102037.

Feng, Yiding, Rad Niazadeh, Amin Saberi. 2019. Linear programming based online policies for real-time assortment of reusable resources. *Available at SSRN 3421227* .

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventory with stochastic demand over finite horizons. *Management Sci.* **40** 999–1020.

Garg, Nikhil, Hamid Nazerzadeh. 2021. Driver surge pricing. *Management Science* .

Gong, Xiao-Yue, Vineet Goyal, Garud Iyengar, David Simchi-Levi, Rajan Udwani, Shuangyu Wang. 2019. Online assortment optimization with reusable resources. *Available at SSRN 3334789* .

Guda, Harish, Upender Subramanian. 2019. Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. *Management Science* **65**(5) 1995–2014.

Hosseini, Mahsa, Joseph Milner, Gonzalo Romero. 2021. Dynamic relocations in car-sharing networks. *Available at SSRN* .

Jasin, S. 2014. Reoptimization and self-adjusting price control for network revenue management. *Oper. Res.* **62** 1168–1178.

Kanoria, Yash, Pengyu Qian. 2019. Near optimal control of a ride-hailing platform via mirror backpressure. *arXiv preprint arXiv:1903.02764* .

Kim, Jeunghyun, Ramandeep S Randhawa. 2018. The value of dynamic pricing in large queueing systems. *Operations Research* **66**(2) 409–425.

Kim, Jeunghyun, Ramandeep S Randhawa, Amy R Ward. 2018. Dynamic scheduling in a many-server, multiclass system: The role of customer impatience in large systems. *Manufacturing & Service Operations Management* **20**(2) 285–301.

Lei, Yanzhe, Stefanus Jasin. 2020. Real-time dynamic pricing for revenue management with reusable resources, advance reservation, and deterministic service time requirements. *Operations Research* .

Lei, Yanzhe, Stefanus Jasin, Amitabh Sinha. 2018. Joint dynamic pricing and order fulfillment for e-commerce retailers. *Manufacturing & Service Operations Management* **20**(2) 269–284.

Levi, R, A Radovanovic. 2010. Provably near-optimal lp-based policies for revenue management in systems with reusable resources. *Oper. Res.* **58**(2) 503–507.

Lim, Adrian. 2017. Taxi companies get green light to implement dynamic pricing system. *The Straits Times (March 17, 2017)* .

New York City. 2020. Taxi & limousine commission trip record data. Data retrieved from `https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page`.

Owen, Z, D Simchi-Levi. 2018. Price and assortment optimization for reusable resources. Accessed at `https://papers.ssrn.com/abstract=3070625`, August 15th, 2019.

Ozkan, Erhun. 2018. Joint pricing and matching in ride-sharing systems. *Available at SSRN 3217642* .

Özkan, Erhun, Amy R Ward. 2020. Dynamic matching for real-time ride sharing. *Stochastic Systems* **10**(1) 29–70.

Rusmevichientong, Paat, Mika Sumida, Huseyin Topaloglu. 2020. Dynamic assortment optimization for reusable products with random usage durations. *Management Science* .

Taylor, Terry A. 2018. On-demand service platforms. *Manufacturing & Service Operations Management* **20**(4) 704–720.

Varma, Sushil Mahavir, Pornpawee Bumpensanti, Siva Theja Maguluri, He Wang. 2022. Dynamic pricing and matching for two-sided queues. *Operations Research* .

Vera, Alberto, Alessandro Arlotto, Itai Gurvich, Eli Levin. 2020. Dynamic resource allocation: The geometry and robustness of constant regret. Tech. rep., Working paper.

Vera, Alberto, Siddhartha Banerjee, Itai Gurvich. 2021. Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research* .

Wang, Yining, He Wang. 2022. Constant regret resolving heuristics for price-based revenue management. *Operations Research* .

Yan, Chiwei, Helin Zhu, Nikita Korolko, Dawn Woodard. 2019. Dynamic pricing and matching in ride-hailing platforms. *Naval Research Logistics (NRL)* .

Yao, Huaxiu, Fei Wu, Jintao Ke, Xianfeng Tang, Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye, Zhenhui Li. 2018. Deep multi-view spatial-temporal network for taxi demand prediction. *Thirty-Second AAAI Conference on Artificial Intelligence*.

Yong, Celement. 2022. More drivers wanted as taxi and private-hire ridership rebounds. *The Straits Times (January 15, 2022)* .

This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

EC.1. Proof of Lemmas

EC.1.1. A Technical Lemma

We first state and prove a technical lemma that will be used repeatedly through the analysis, which essentially is a large deviation bound on the absolute magnitude of cumulative demand variability along arcs across consecutive time periods.

LEMMA EC.1. *For any $1 \le t_1 < t_2 \le T$, any arc $i \to j$, and any $y \ge 0$ and $r \in (0,1]$, it holds that*

$$\mathbf{P}\left(\max_{t_1\le t\le t_2}\left|\sum_{s=t_1}^{t}\delta_{s,ij}\right| > y\right) \le 2\exp\left(\sum_{s=t_1}^{t_2}\min\{\lambda_{s,ij}, 1-\lambda_{s,ij}\}\cdot r^2 - ry\right) \quad \text{(EC.1)}$$

$$\le 2\exp((t_2-t_1)r^2 - ry) \quad \text{(EC.2)}$$

Proof of Lemma EC.1. The inequality (EC.2) follows directly from inequality (EC.1) since $\lambda_{s,ij} \in [0,1]$. Next we prove inequality (EC.1). For any $r \in (0,1]$, the following holds

$$\begin{aligned}
\mathbf{P}\left(\max_{t_1\le t\le t_2}\left|\sum_{s=t_1}^{t}\delta_{s,ij}\right| > y\right) &= \mathbf{P}\left(\max_{t_1\le t\le t_2}\exp(r\left|\sum_{s=t_1}^{t}\delta_{s,ij}\right|) > e^{ry}\right) \\
&\le \mathbf{P}\left(\max_{t_1\le t\le t_2}\exp(r\sum_{s=t_1}^{t}\delta_{s,ij}) > e^{ry}\right) + \mathbf{P}\left(\max_{t_1\le t\le t_2}\exp(-r\sum_{s=t_1}^{t}\delta_{s,ij}) > e^{ry}\right) \\
&\le \mathbf{E}\left[\exp(r\sum_{s=t_1}^{t_2}\delta_{s,ij} - ry) + \exp(-r\sum_{s=t_1}^{t_2}\delta_{s,ij} - ry)\right] \qquad \text{(EC.3)}
\end{aligned}$$

where the first inequality follows by the union bound and the second inequality follows by Doob's martingale inequality. Note that, by elementary algebra, we have

$$\mathbf{E}\exp(r\delta_{s,ij}) = (\lambda_{s,ij}e^r + 1 - \lambda_{s,ij})\exp(-r\lambda_{s,ij}) \le \exp((e^r - 1 - r)\lambda_{s,ij}) \le \exp(r^2\lambda_{s,ij}) \quad \text{(EC.4)}$$

where the equality follows since $\lambda_{s,ij} + \delta_{s,ij}$ is a binary random variable that equals 1 with probability $\lambda_{s,ij}$, the first inequality follows since $e^x \ge 1 + x$, the second inequality follows since $e^r - 1 - r \le r^2$ for all $r \in [-1,1]$. On the other hand, we obtain an alternative upper bound in the following way:

$$\begin{aligned}
\mathbf{E}\exp(r\delta_{s,ij}) &= (\lambda_{s,ij}e^r + 1 - \lambda_{s,ij})\exp(-r\lambda_{s,ij}) = (1-(1-\lambda_{s,ij})(1-e^{-r}))\exp(r(1-\lambda_{s,ij})) \\
&\le \exp((e^{-r} - 1 + r)(1-\lambda_{s,ij})) \le \exp(r^2(1-\lambda_{s,ij})) \qquad \text{(EC.5)}
\end{aligned}$$

where the second equality follows from elementary algebra, the first inequality holds since $1 - x \le e^{-x}$. We can apply similar techniques to obtain two more bounds as follows:

$$\mathbf{E}\exp(-r\delta_{s,ij}) = (\lambda_{s,ij}e^{-r} + 1 - \lambda_{s,ij})\exp(r\lambda_{s,ij}) \le \exp(e^{-r} - 1 + r)\lambda_{s,ij}) \le \exp(r^2\lambda_{s,ij}), \quad \text{(EC.6)}$$

$$\begin{aligned}
\mathbf{E}\exp(-r\delta_{s,ij}) &= (\lambda_{s,ij}e^{-r} + 1 - \lambda_{s,ij})\exp(r\lambda_{s,ij}) = (1+(1-\lambda_{s,ij})(e^r - 1))\exp(-r(1-\lambda_{s,ij})) \\
&\le \exp((e^{-r} - 1 - r)(1-\lambda_{s,ij})) \le \exp(r^2(1-\lambda_{s,ij})) \qquad \text{(EC.7)}
\end{aligned}$$

Inequality (EC.1) then follows by combining inequalities (EC.3) - (EC.7). $\square$

EC.1.2. Proof of Lemma 1

Take any control $\pi \in \Pi$ that is feasible to **SDP**. Define $\lambda^*_{t,ij} := \mathbf{E}^\pi[\lambda_{t,ij}(p^\pi_{t,ij})]$. Note that $\{\lambda^*_{t,ij}\}_{t,ij}$ is feasible to $\mathbf{DCP}(0)$. Indeed, (3) implies that under π, $\lambda_{t,ij}(p^\pi_{t,ij}) \in [0,1]$; thus, $\lambda^*_{t,ij} = \mathbf{E}^\pi[\lambda_{t,ij}(p^\pi_{t,ij})] \in [0,1]$. Moreover, by (2), the following holds for all i, t,

$$C_i - \sum_{j=1}^N \sum_{s=1}^t \lambda^*_{s,ij} + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji})^+} \lambda^*_{s,ji} = \mathbf{E}^\pi \left[C_i - \sum_{j=1}^N \sum_{s=1}^t D_{s,ij}(p^\pi_{s,ij}) + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji})^+} D_{s,ji}(p^\pi_{s,ji}) \right] \geq 0,$$

We have just shown that $\{\lambda^*_{t,ij}\}_{t,ij}$ is feasible to $\mathbf{DCP}(0)$. Finally, note that

$$\begin{aligned}
\mathbf{E}^\pi \left[p^\pi_{t,ij} D_{t,ij}(p^\pi_{t,ij}) \right] &= \mathbf{E}^\pi \left\{ \mathbf{E}^\pi \left[p^\pi_{t,ij} D_{t,ij}(p^\pi_{t,ij}) \middle| \mathcal{F}_t \right] \right\} \\
&= \mathbf{E}^\pi \left[p^\pi_{t,ij} \lambda_{t,ij}(p^\pi_{t,ij}) \right] = \mathbf{E}^\pi \left[r_{t,ij}(\lambda_{t,ij}(p^\pi_{t,ij})) \right] \\
&\leq r_{t,ij} \left(\mathbf{E}^\pi \left[\lambda_{t,ij}(p^\pi_{t,ij}) \right] \right) = r_{t,ij}(\lambda^*_{t,ij}),
\end{aligned}$$

where the inequality follows by the Jensen's inequality and A2. We then conclude that $\mathcal{J}(0) \geq \mathcal{J}^\star$ since $\mathcal{J}(0) \geq \sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N r_{t,ij}(\lambda^*_{t,ij}) \geq \mathbf{E}^\pi[\sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N p^\pi_{t,ij} D_{t,ij}(p^\pi_{t,ij})]$. $\square$

EC.1.3. Proof of Lemma 2

Note that $\mathbf{DCP(0)}$ is a feasible convex program and it satisfies the Slater's condition. Moreover, there exists $\bar{\zeta} > 0$ such that $\mathbf{DCP(\zeta)}$ is also feasible for all $\zeta \leq \bar{\zeta}$. Let Ψ'_0 denote the largest dual variable of $\mathbf{DCP(0)}$ for all the constraints in (6). Then, by strong duality, $\mathcal{J}^{Det} - \mathcal{J}(\zeta) = \mathcal{J}(0) - \mathcal{J}(\zeta) \leq 2\Psi'_0 N^2 T \zeta$ holds for all $\zeta \leq \bar{\zeta}$. Setting $\Psi_0 = 2\Psi'_0 N^2$ concludes the proof. $\square$.

EC.1.4. Proof of Lemma 3

We prove by induction. For the base case, $\mathbb{H}_1$ holds since, by definition, $C^n_{1,i} = nC_{1,i} \geq nN \geq N$; so $p^\pi_{1,ij} = p^n_{1,ij}(\lambda^{n,\zeta,D}_{1,ij} - \epsilon)$ by the definition of SPC.

For the inductive step, suppose that $\mathbb{H}_s$ holds for all $s \leq t$. We now show that $\mathbb{H}_{t+1}$ holds as well. Note that the number of supply in node i cannot increase before period $\min_j \tau^n_{ji}$, since no supply units will flow in to node i before that period. Therefore, we need to argue for the case when $t < \min_j \tau^n_{ji}$ and when $t \geq \min_j \tau^n_{ji}$ separately. If $t < \min_j \tau^n_{ji}$, then we have

$$\begin{aligned}
C^n_{t+1,i} &= C^n_{1,i} - \sum_{j=1}^N \sum_{s=1}^t D^n_{s,ij} = C^n_{1,i} - \sum_{j=1}^N \sum_{s=1}^t (\lambda^{n,\zeta,D}_{s,ij} + D^n_{s,ij} - \lambda^{n,\zeta,D}_{s,ij}) \\
&\geq \sum_{j=1}^N \sum_{s=1}^{\min_j \tau^n_{ji}} \lambda^{n,\zeta,D}_{s,ij} - \sum_{j=1}^N \sum_{s=1}^t \lambda^{n,\zeta,D}_{s,ij} - \sum_{j=1}^N \sum_{s=1}^t (D^n_{s,ij} - \lambda^{n,\zeta,D}_{s,ij}) \\
&= \sum_{j=1}^N \left(\sum_{s=t+1}^{\min_j \tau^n_{ji}} \lambda^{n,\zeta,D}_{s,ij} + \sum_{s=1}^t (\lambda^{n,\zeta,D}_{s,ij} - \lambda^n_{s,ij}(p^\pi_{s,ij})) - \sum_{s=1}^t \delta^n_{s,ij} \right)
\end{aligned}$$

$$\geq N\left(\left(\min_j \tau_{ji}^n - t\right)\cdot\zeta + t\cdot\epsilon\right) - \sum_{j=1}^N \left|\sum_{s=1}^t \delta_{s,ij}^n\right|$$
$$\geq N\epsilon\underline{\tau}^n - N(\epsilon\underline{\tau}^n - 1) = N \tag{EC.8}$$

where the first inequality holds by inequality (5) at period $\min_j \tau_{ji}^n$; the second inequality holds by the inductive assumption $\mathbb{H}_s$ and the fact that $\lambda_{s,ij}^{n,\zeta,D} \geq \zeta$ (inequality (6)), the third equality holds by the definition of $\mathcal{S}_{ij}$ and $\zeta = \epsilon$. On the other hand, if $t \geq \min_j \tau_{ji}^n$, we know that

$$\begin{aligned}
C_{t+1,i}^n &= C_{1,i}^n - \sum_{j=1}^N\sum_{s=1}^t D_{s,ij}^n + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji}^n)^+} D_{s,ji}^n \\
&= C_{1,i}^n - \sum_{j=1}^N\sum_{s=1}^t (\lambda_{s,ij}^{n,\zeta,D} + D_{s,ij}^n - \lambda_{s,ij}^{n,\zeta,D}) + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji}^n)^+} (\lambda_{s,ji}^{n,\zeta,D} + D_{s,ji}^n - \lambda_{s,ji}^{n,\zeta,D}) \\
&\geq -\sum_{j=1}^N\sum_{s=1}^t (D_{s,ij}^n - \lambda_{s,ij}^{n,\zeta,D}) + \sum_{j=1}^N \sum_{s=1}^{(t-\tau_{ji}^n)^+} (D_{s,ji}^n - \lambda_{s,ji}^{n,\zeta,D}) \\
&\geq \sum_{j=1}^N \left[\sum_{s=1}^t (\lambda_{s,ij}^{n,\zeta,D} - \lambda_{s,ij}^n(p_{s,ij}^\pi)) - \sum_{s=1}^{(t-\tau_{ji}^n)^+} (\lambda_{s,ji}^{n,\zeta,D} - \lambda_{s,ji}^n(p_{s,ji}^\pi))\right] - \sum_{j=1}^N \left(\left|\sum_{s=1}^t \delta_{s,ij}^n\right| + \left|\sum_{s=1}^{(t-\tau_{ji}^n)^+} \delta_{s,ji}^n\right|\right) \\
&\geq \epsilon\left(Nt - \sum_{j=1}^N (t-\tau_{ji}^n)^+\right) - N(\epsilon\underline{\tau}^n - 1) \\
&\geq N\epsilon\underline{\tau}^n - N(\epsilon\underline{\tau}^n - 1) = N
\end{aligned} \tag{EC.9}$$

where the first inequality follows since $\{\lambda_{t,ij}^{n,\zeta,D}\}_{t,ij}$ is feasible to $\mathbf{DCP}(n,\zeta)$, the second inequality follows since $D_{s,ij}^n = \lambda_{s,ij}^n(p_{s,ij}^\pi) + \delta_{s,ij}^n$, the third inequality follows by the inductive assumption and the definition of $\mathcal{S}_{ij}$, and the fourth inequality holds since $t \geq \min_j \tau_{ji}^n \geq \underline{\tau}^n$. Moreover, inequalities (EC.8) and (EC.9) implies that $p_{t+1,ij}^\pi = p_{t+1,ij}^n(\lambda_{t+1,ij}^{n,\zeta,D} - \epsilon)$ by the definition of SPC. Hence, $\mathbb{H}_{t+1}$ holds which completes the inductive step.

We continue to prove the probability bound. By Lemma EC.1 and letting $r = (\epsilon\underline{\tau}^n - 1)/(4T^n) = \sqrt{\log(n\underline{\tau})/(nT)} \in (0,1]$, we have:

$$\begin{aligned}
\mathbf{P}(\mathcal{S}^c) &\leq \sum_{i=1}^N\sum_{j=1}^N \mathbf{P}(\mathcal{S}_{ij}^c) \leq 2N^2\exp\left(T^n r^2 - r\frac{\epsilon\underline{\tau}^n - 1}{2}\right) \\
&= 2N^2\exp\left(-\frac{(\epsilon\underline{\tau}^n - 1)^2}{16T^n}\right) = 2N^2\exp(-\log(n\underline{\tau})) = \frac{2N^2}{n\underline{\tau}}. \quad \square
\end{aligned}$$

EC.1.5. Proof of Lemma 4

We first show that, for all sample paths on $\mathcal{A}$, it holds that $\hat{\lambda}_{t,ij}^\pi = \lambda_{t,ij}^{n,\zeta,D} - \epsilon - u_{t,ij}\cdot\bar{\delta}_{ij}^{\kappa_{ij}(t)-1} \in (0,1)$. In other words, the targeted demand rate after adjustment is still in the interior of feasible region for demand rate. Fix i, j and k. Consider two cases. For any $t \in \mathcal{T}_{ij}^k$, if $\bar{\delta}_{ij}^{k-1} \geq 0$, then we have

$$1 > \lambda_{t,ij}^{n,\zeta,D} - \epsilon \geq \lambda_{t,ij}^{n,\zeta,D} - \epsilon - u_{t,ij}\cdot\bar{\delta}_{ij}^{k-1}$$

$$
\begin{aligned}
&= \frac{\lambda_{t,ij}^{n,\zeta,D}}{2} - \epsilon + \frac{\lambda_{t,ij}^{n,\zeta,D}}{2} - \frac{\lambda_{t,ij}^{n,\zeta,D}(1-\lambda_{t,ij}^{n,\zeta,D})}{\sum_{v\in\mathcal{T}_{ij}^k}\lambda_{v,ij}^{n,\zeta,D}(1-\lambda_{v,ij}^{n,\zeta,D})}\cdot\bar{\delta}_{ij}^{k-1} \\
&\geq \frac{\zeta}{2} - \epsilon + \lambda_{t,ij}^{n,\zeta,D}\cdot\left(\frac{1}{2}-\frac{\eta}{b}\right) > 0,
\end{aligned}
$$

where the second inequality holds since $\bar{\delta}_{ij}^{k-1}\geq 0$, the third inequality holds by the definition of batch $\mathcal{T}_{ij}^k$ and event $\mathcal{A}$, and the last inequality holds by condition **C2**. On the other hand, if $\bar{\delta}_{ij}^{k-1}<0$, then we have

$$
\begin{aligned}
0 &\leq \lambda_{t,ij}^{n,\zeta,D} - \epsilon \leq \lambda_{t,ij}^{n,\zeta,D} - \epsilon - u_{t,ij}\cdot\bar{\delta}_{ij}^{k-1} = \lambda_{t,ij}^{n,\zeta,D} - \epsilon - \frac{\lambda_{t,ij}^{n,\zeta,D}(1-\lambda_{t,ij}^{n,\zeta,D})}{\sum_{v\in\mathcal{T}_{ij}^k}\lambda_{v,ij}^{n,\zeta,D}(1-\lambda_{v,ij}^{n,\zeta,D})}\cdot\bar{\delta}_{ij}^{k-1} \\
&\leq \lambda_{t,ij}^{n,\zeta,D} - \epsilon + \left(1-\lambda_{t,ij}^{n,\zeta,D}\right)\cdot\frac{\eta}{b} < \lambda_{t,ij}^{n,\zeta,D} - \epsilon + 1 - \lambda_{t,ij}^{n,\zeta,D} < 1
\end{aligned}
$$

where the second inequality holds since $\bar{\delta}_{ij}^{k-1}<0$, the third inequality holds by the definition of $\mathcal{A}$, definition of batch $\mathcal{T}_{ij}^k$, and condition **C2**, and the fourth inequality holds since $\eta/b<1/2<1$. Therefore, we can conclude that $\lambda_{t,ij}^{\zeta,D}-\epsilon-u_{s,ij}\cdot\bar{\delta}_{ij}^{\kappa_{ij}(t)}\in(0,1)$ holds on event $\mathcal{A}$.

We now proceed to prove condition $\mathbb{H}_t$ by induction. When $t=1$, $C_{1,i}^n = nC_{1,i}\geq N$ by definition, which also implies that $p_{1,ij}^{\pi} = p_{1,ij}^n(\hat{\lambda}_{t,ij}) = p_{1,ij}^n(\lambda_{t,ij}^{n,\zeta,D}-\epsilon)$ by the definition of ABC. This completes the inductional basis. Suppose that $\mathbb{H}_s$ holds for all $s\leq t$, we now show that $\mathbb{H}_{t+1}$ holds as well. Note that the following holds for all sample paths on $\mathcal{A}$:

$$
\begin{aligned}
C_{t+1,i}^n &\geq N\epsilon\underline{\tau}^n + \sum_{j=1}^N\left(\sum_{s=1}^{(t-\tau_{ji}^n)^+}\delta_{s,ji}^n + \sum_{s=1}^{(t-\tau_{ji}^n)^+}u_{s,ji}\cdot\bar{\delta}_{ji}^{\kappa_{ji}(s)-1}\right) - \sum_{j=1}^N\left(\sum_{s=1}^t\delta_{s,ij}^n + \sum_{s=1}^t u_{s,ij}\cdot\bar{\delta}_{ij}^{\kappa_{ij}(s)-1}\right) \\
&\geq N\epsilon\underline{\tau}^n + \sum_{j=1}^N\sum_{s\in\mathcal{T}_{ji}^{\kappa_{ji}((t-\tau_{ji}^n)^+)}}\delta_{s,ji}^n + \sum_{j=1}^N\left(1-\sum_{s=\min\left\{v\in\mathcal{T}_{ji}^{\kappa_{ji}((t-\tau_{ji}^n)^+)}\right\}}^{(t-\tau_{ji}^n)^+}u_{s,ji}\right)\cdot\bar{\delta}_{ji}^{\kappa_{ji}((t-\tau_{ji}^n)^+)-1} \\
&\quad -\sum_{j=1}^N\sum_{s\in\mathcal{T}_{ij}^{\kappa_{ij}(t)}}\delta_{s,ij}^n + \sum_{j=1}^N\left(1-\sum_{s=\min\left\{v\in\mathcal{T}_{ij}^{\kappa_{ij}(t)}\right\}}^{t}u_{s,ij}\right)\cdot\bar{\delta}_{ij}^{\kappa_{ij}(t)-1}
\end{aligned}
\tag{EC.10}
$$

where the first inequality follows by the same argument as in the proof of (EC.8) and (EC.9) and the inductional hypothesis, and the last inequality follows by the definition of $\bar{\delta}_{ij}^k$ and the fact that $\sum_{t\in\mathcal{T}_{ij}^k}u_{t,ij}=1$ for all i, j and k. Note that the absolute value of the second until the fifth terms after the last equality in (EC.10) are all bounded from above by $\sum_{j=1}^N\eta = N\eta$ on event $\mathcal{A}$. Therefore, we can bound

$$
C_{t+1,i}^n \geq N\epsilon\underline{\tau}^n - 4N\eta = N \tag{EC.11}
$$

where the equality follows by the definition of η. Since the supply level is positive at node i, by Step 2.b of ABC, we know that $p_{t+1,ij}^{\pi} = p_{t+1,ij}^n(\hat{\lambda}_{t+1,ij})$. This completes the inductional step.

We still need to prove the probability bound. Note that, for each arc $i \to j$, the maximum number of batches K_{ij} is at most T^n/b since expected demand rate in one period is at most one. For each $i \to j$ and any $K_{ij}+1 \le k \le T^n/b$, we further define $\mathcal{T}_{ij}^k = \emptyset$. Then, applying union bound we have

$$\begin{aligned}
\mathbf{P}(\mathcal{A}^c) &= \mathbf{P}\left(\cup_{i=1}^N \cup_{j=1}^N \cup_{k=1}^{K_{ij}} \mathcal{A}_{ijk}^c\right) \le \mathbf{P}\left(\cup_{i=1}^N \cup_{j=1}^N \cup_{k=1}^{T^n/b} \mathcal{A}_{ijk}^c\right) \le \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^{T^n/b} \mathbf{P}(\mathcal{A}_{ijk}^c) \\
&\le \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^{T^n/b} \mathbf{P}\left\{\max_{t \in \mathcal{T}_{ij}^k} \left| \sum_{s=\min\{v \in \mathcal{T}_{ij}^k\}}^{t} \delta_{s,ij}^n \right| \ge \eta \right\} \\
&\le \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^{T^n/b} 2\exp\left(\sum_{s \in \mathcal{T}_{ij}^k} \min\{\lambda_{s,ij}, 1-\lambda_{s,ij}\} \cdot r^2 - \eta \cdot r\right) \\
&\le \frac{2N^2T^n}{b} \exp\left(4(b+1)r^2 - \eta \cdot r\right) \le \frac{2N^2T^n}{b} \exp\left(-\frac{(\epsilon \underline{T}^n - 1)^2}{256(b+1)}\right) \\
&\le \frac{2N^2}{bT^n}
\end{aligned}$$

where the fourth inequality holds by (EC.1) in Lemma EC.1, the third to the last inequality holds by the the fact that $\min\{x, 1-x\} \le 2x(1-x)$ holds for any $x \in (0,1)$ and the definition of $\mathcal{T}_{ij}^k$, the second to the last inequality holds by setting $r = \eta/(8(b+1)) \in (0,1)$ (the inclusion follows by **C2**), the last inequality follows by the definition of ϵ. $\square$

EC.2. Proof of Theorem 1: Lower Bound on Static Policy

Fix n. Note that the objective function of the deterministic problem $\mathbf{DCP}(n,0)$ equals $\sum_{t=1}^{nT}[p_{t,12}(1-p_{t,12}) + p_{t,21}(1-p_{t,21})]$, and it achieves the unconstrained optimum at $p_{t,12} = p_{t,21} = 0.5$ for all t with an optimal revenue of $nT/2$. Moreover, it can be easily verified that this set of prices satisfies the constraints of $\mathbf{DCP}(n,0)$: Specifically, the capacity at node i decreases at a rate of 0.5 units per time unit for a total of $n\tau_{ij} = 2n$ time units, reaching a capacity level of 0 at the end of period $n\tau_{ij}$; afterwards, the incoming flow and outgoing flow at node i balances out so the capacity is kept at 0 at each node. Take any $\pi \in \Pi$, then,

$$\begin{aligned}
nT\mathcal{L}^\pi &= \mathcal{J}(n,0) - \mathbf{E}^\pi[R^\pi(n)] = \frac{nT}{2} - \mathbf{E}^\pi[R^\pi(n)] \\
&= \frac{n\tau}{2} - \mathbf{E}^\pi\left[\sum_{t=1}^{n\tau}\sum_{i,j} D_{t,ij}(p_{t,ij}^\pi)p_{t,ij}^\pi\right] + \frac{n(T-\tau)}{2} - \mathbf{E}^\pi\left[\sum_{t=n\tau+1}^{nT}\sum_{i,j} D_{t,ij}(p_{t,ij}^\pi)p_{t,ij}^\pi\right] \\
&\ge \frac{n\tau}{2} - \mathbf{E}^\pi\left[\sum_{t=1}^{n\tau}\sum_{i,j} D_{t,ij}(p_{t,ij}^\pi)p_{t,ij}^\pi\right],
\end{aligned} \tag{EC.12}$$

where the inequality follows since

$$\mathbf{E}^\pi\left[\sum_{t=n\tau+1}^{nT}\sum_{i,j} D_{t,ij}(p_{t,ij}^\pi)p_{t,ij}^\pi\right] = \sum_{t=n\tau+1}^{nT} \mathbf{E}^\pi\left[\mathbf{E}^\pi\left[\sum_{i,j} D_{t,ij}(p_{t,ij}^\pi)p_{t,ij}^\pi \,\middle|\, \mathcal{F}_t\right]\right]$$

$$= \sum_{t=n\tau+1}^{nT} \mathbf{E}^\pi \left[\sum_{i,j} \lambda_{t,ij}(p^\pi_{t,ij}) p^\pi_{t,ij} \right] \leq \sum_{t=n\tau+1}^{nT} \sum_{i,j} \lambda_{t,ij}(\mathbf{E}^\pi[p^\pi_{t,ij}]) \mathbf{E}^\pi[p^\pi_{t,ij}]$$
$$\leq \sum_{t=n\tau+1}^{nT} \max_{x_{t,12}, x_{t,21}} \sum_{i,j} \lambda_{t,ij}(x_{t,ij}) x_{t,ij} = \frac{n(T-\tau)}{2},$$

where the first inequality follows by Jensen's inequality and the second inequality follows by principle of optimization. Finally, note that since the resource cannot reach the destination within $n\tau$ days, this means that (EC.12) equals the revenue gap between the deterministic upper bound and the revenue of a static pricing policy for a classic network revenue management problem with two resources each with capacity $n\tau/2 = n$, two products each with an identical and independent demand function $\lambda(p) = 1 - p$, a capacity consumption matrix $[1\ 0; 0\ 1]$, and a planning horizon of $n\tau = 2n$ periods, which is known in the literature to have a revenue loss in the order of $\sqrt{n}$ (see Remark 2 in Jasin (2014)), i.e., there exists some constant $\check{\psi}$ independent of n such that

$$\frac{n\tau}{2} - \mathbf{E}^\pi \left[\sum_{t=1}^{n\tau} \sum_{i,j} D_{t,ij}(p^\pi_{t,ij}) p^\pi_{t,ij} \right] \leq \check{\psi}\sqrt{n\tau}.$$

The stated result follows by letting $\tilde{\Psi}_0 = \check{\psi}\sqrt{\tau}/T$. □

EC.3. Additional Details to the Numerical Studies

In this section, we report additional details of the numerical studies. All the experiments are implemented on a Windows desktop with Intel(R) Xeon(R) W-2145 CPU and 64 GB RAM.

EC.3.1. Synthetic Data

The travel times among different nodes and the initial server distribution are given as follows.

$$[\tau_{ij}] = \begin{bmatrix} 0 & 7 & 10 & 8 & 2 \\ 8 & 0 & 9 & 4 & 10 \\ 9 & 11 & 0 & 6 & 9 \\ 7 & 5 & 6 & 0 & 9 \\ 2 & 9 & 10 & 8 & 0 \end{bmatrix}, \quad C = [5, 3, 4, 6, 5]^\top.$$

Table EC.1 reports the performance of different pricing controls (as compared with $\mathcal{J}(\zeta)$) when stationary optimal deterministic solution is used as the baseline control. In particular, the only difference between SPC (resp. ABC) and S-SPC (resp. S-ABC) is that the second heuristic uses $\{\lambda^{\zeta,SD}_{t,ij}\}_{t,ij}$ in Step 1 as the baseline control, where $\lambda^{\zeta,SD}_{t,ij} \equiv \lambda^{\zeta,SD,*}_{ij}$ for all t and $\lambda^{\zeta,SD,*}_{ij}$ is the solution to the following optimization problem:

$$\textbf{(S-DCP}(\zeta)\textbf{)} \quad \mathcal{J}^{\text{STAT}}(\zeta) := \max_{(\lambda_{ij})_{(i,j)}} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} r_{t,ij}(\lambda_{ij})$$
$$s.t. \quad C_i - \sum_{j=1}^{N} t \cdot \lambda_{ij} + \sum_{j=1}^{N} (t - \tau_{ji})^+ \lambda_{ji} \geq 0, \forall i, t$$
$$\zeta \leq \lambda_{ij} \leq 1 - \zeta, \forall i, j$$

As a reference, $(\mathcal{J}(\zeta) - \mathcal{J}^{\text{STAT}}(\zeta))/\mathcal{J}(\zeta) = 6.93\%$. As we can observe, the performance of S-ABC is worse than that of SPC, which suggests the necessity of using non-stationary baseline control even if the prices are adjusted dynamically.

n	S-SPC Loss $	S-SPC Loss %	S-SPC Rev. $	S-ABC Loss $	S-ABC Loss %	S-ABC Rev. $	S-ABC Rev. Incr. %
40	14.30	12.08	124879	9.41	7.95	130748	4.70
80	12.58	10.63	253890	8.87	7.50	262777	3.50
120	11.82	9.98	383568	8.76	7.40	394564	2.87
160	11.40	9.63	513411	8.68	7.33	526483	2.55
200	11.16	9.43	643207	8.65	7.31	658286	2.34
240	10.95	9.25	773407	8.55	7.23	790648	2.23
280	10.83	9.15	903266	8.54	7.21	922565	2.14
320	10.66	9.01	1033951	8.48	7.17	1054859	2.02
360	10.54	8.91	1164494	8.45	7.13	1187129	1.94
400	10.49	8.86	1294496	8.40	7.10	1319549	1.94

Table EC.1 Performance of proposed pricing controls under varying n using stationary optimal deterministic solution as baseline

We further report the experiment results under different demand models: Table EC.2 contains the results under exponential demand models (where $\lambda_{t,ij} = \mu_{ij} \cdot \exp(-\beta_{ij} \cdot p_{t,ij})$); Table EC.3 contains the results under logit demand models (where $\lambda_{t,ij}(p_{t,ij}) = \mu_{ij} \cdot \frac{\exp(\alpha - \beta_{ij} \cdot p_{t,ij})}{1+\exp(\alpha - \beta_{ij} \cdot p_{t,ij})}$). The parameters are set in the same way as in the case with linear demand models. The observations from these two tables are qualitatively similar to those drawn under the linear demand models.

n	SPC Loss $	SPC Loss %	SPC Rev. $	SPC Admit	ABC Loss $	ABC Loss %	ABC Rev. $	ABC Rev. Incr. %	ABC Admit Incr. %	ABC Avg Price Incr. %
40	4.86	5.41	101974	1950	2.58	2.88	104706	2.68	1.61	1.05
80	3.66	4.08	206821	3948	1.72	1.91	211493	2.26	1.51	0.74
120	3.11	3.46	312239	5957	1.20	1.34	319095	2.20	1.54	0.65
160	2.78	3.09	417889	7972	1.02	1.14	426312	2.02	1.44	0.56
200	2.53	2.82	523853	9981	0.88	0.98	533775	1.89	1.23	0.66
240	2.31	2.57	630184	11996	0.74	0.83	641484	1.79	1.15	0.63
280	2.29	2.55	735412	13996	0.73	0.81	748529	1.78	0.94	0.84
320	2.04	2.28	842825	16023	0.71	0.79	855643	1.52	0.82	0.69
360	1.78	1.98	951051	18046	0.63	0.70	963443	1.30	0.62	0.68
400	1.65	1.83	1058311	20070	0.53	0.59	1071652	1.26	0.65	0.60

Table EC.2 Performance of proposed pricing controls under varying n with exponential demand function

	SPC				ABC					
n	Loss $	Loss %	Rev. $	Admit	Loss $	Loss %	Rev. $	Rev. Incr. %	Admit Incr. %	Avg Price Incr. %
40	12.84	3.70	400475	3293	6.74	1.95	407786	1.83	1.25	0.57
80	10.25	2.96	807153	6631	4.53	1.31	820887	1.70	1.21	0.49
120	8.95	2.58	1215424	9980	3.72	1.07	1234247	1.55	1.11	0.43
160	8.32	2.40	1623596	13331	3.02	0.87	1648998	1.56	1.10	0.46
200	7.82	2.26	2032477	16686	2.80	0.81	2062588	1.48	1.09	0.39
240	6.62	1.91	2447594	20036	2.31	0.67	2478602	1.27	0.84	0.42
280	6.15	1.77	2859492	23404	2.07	0.60	2893758	1.20	0.78	0.42
320	6.10	1.76	3268465	26751	2.00	0.58	3307832	1.20	0.66	0.54
360	5.83	1.68	3679970	30117	1.96	0.56	3721759	1.14	0.64	0.49
400	5.57	1.61	4091908	33487	1.62	0.47	4139397	1.16	0.61	0.55

Table EC.3 Performance of proposed pricing controls under varying n with logit demand function

At last, we test the performance of our proposed controls by varying the length of planning horizon. Specifically, we fix the scaling parameter $n = 100$ and set $T = \{30, 40, 50, 60, 70, 80\}$. Table EC.4 reports the numerical results. We find that, as T increases, both SPC and ABC incur higher revenue loss, but the loss of ABC grows slower than that of SPC. In other words, the performance of ABC is more robust under different length of planning horizon than that of SPC.

		SPC			ABC			
T	$\mathcal{J}(1,0)$	Loss $	Loss %	Rev. $	Loss $	Loss %	Rev. $	Rev. Incr. %
30	3551	5.81	4.91	3377	3.33	2.82	3451	2.20
40	4639	6.28	5.41	4388	3.35	2.89	4505	2.67
50	5720	6.70	5.85	5385	3.47	3.03	5547	3.00
60	6807	6.82	6.01	6398	3.52	3.10	6595	3.09
70	7891	7.08	6.28	7396	3.56	3.16	7642	3.33
80	8976	7.33	6.53	8389	3.55	3.16	8692	3.61

Table EC.4 Performance of pricing controls under varying T

EC.3.2. New York Taxi Dataset

Table EC.5 summarize how we construct regions by grouping adjacent regions defined in New York City (2020). Roughly speaking, the bigger the difference of the region indexes is, the farther two regions are apart. Table EC.6 provides summary statistics for the travel time from region 1 to all the other regions. In general, the travel time is proportional to the distance.

We now give a formal definition for the deterministic relaxation problem used in Section 6.2 where the optimal static solution is piece-wise stationary. Define U to be a positive integer and assume without loss of generality that T is integral multiply of U. In particular, U should be

Region ID	**Original Zone Number and Names**
1	12-Battery Park, 13-Battery Park City , 261-World Trade Center
2	87-Financial District North, 88-Financial District South , 209-Seaport
3	125-Hudson Sq, 211-SoHo , 231-TriBeCa/Civic Center
4	45-Chinatown, 144-Little Italy/NoLiTa , 148-Lower East Side
5	158-Meatpacking/West Village West, 249-West Village
6	79-East Village, 113-Greenwich Village North , 114-Greenwich Village South
7	4-Alphabet City, 232-Two Bridges/Seward Park
8	68-East Chelsea, 246-West Chelsea/Hudson Yards
9	90-Flatiron, 100-Garment District , 186-Penn Station/Madison Sq West
10	107-Gramercy, 224-Stuy Town/Peter Cooper Village , 234-Union Sq
11	137-Kips Bay, 164-Midtown South , 170-Murray Hill
12	48-Clinton East, 50-Clinton West
13	161-Midtown Center, 163-Midtown North , 230-Times Sq/Theatre District
14	162-Midtown East, 229-Sutton Place/Turtle Bay North , 233-UN/Turtle Bay South
15	142-Lincoln Square East, 143-Lincoln Square West , 140-Lenox Hill East
16	140-Lenox Hill East, 141-Lenox Hill West , 237-Upper East Side South
17	238-Upper West Side North, 239-Upper West Side South
18	236-Upper East Side North, 262-Yorkville East , 263-Yorkville West
19	24-Bloomingdale, 151-Manhattan Valley
20	43-Central Park

Table EC.5 Construction of Region IDs used in the Numerical Experiments

Destination	Median	25%-quantile	75%-quantile	10%-quantile	90%-quantile
2	529.75	508.45	561.43	478.73	587.21
3	502.61	474.71	530.46	451.09	557.15
4	915.61	864.21	966.23	809.54	1022.73
5	638.67	609.86	679.01	585.03	708.49
6	951.55	906.91	1009.68	862.89	1046.54
7	899.40	835.18	964.57	781.50	1027.38
8	929.53	872.28	998.34	819.12	1081.04
9	1300.08	1199.28	1397.63	1112.33	1503.27
10	1213.90	1145.57	1268.94	1075.84	1321.19
11	1452.32	1351.50	1580.42	1260.00	1676.11
12	1506.24	1359.06	1674.97	1259.11	1842.51
13	1778.76	1608.01	1937.79	1491.31	2081.46
14	1412.46	1297.01	1535.53	1168.00	1666.09
15	1692.00	1552.61	1877.09	1402.41	2037.61
16	1747.87	1570.98	1911.89	1364.52	2060.91
17	1853.75	1682.08	2061.70	1493.60	2275.25
18	1992.82	1782.82	2196.30	1610.23	2367.80
19	1971.50	1748.31	2258.25	1502.13	2621.10
20	2055.18	1870.63	2239.57	1681.13	2411.45

Table EC.6 Average Travel Time From Region 1 in Seconds

interpreted as the duration during which demand rate is piece-wise stationary. In our experiment, since one period equals 5 seconds and optimal demand rate is required to be stationary during each 5-minute time window, $U = 5\times 60/5 = 60$. Let $\mathcal{U} = \{\mathcal{U}_k\}$ to be a partition of $[T]$, where $\mathcal{U}_k = \{(k-1)t+1,\dots,kt\}$. Denote by Π^U the class of controls where the prices for all the periods in $\mathcal{U}_k$ are the same, i.e. $\Pi(U) = \{\pi : p^{\pi}_{t,ij} = p_{s,ij}, \forall s,t\in\mathcal{U}_k, \forall k\}$. In other words, $\Pi(U)\subset\Pi$ is the class of admissible controls where the price trajectories are piece-wise stationary (in other words, we restrict our attention to the class of pricing controls under which the prices cannot be adjusted more frequently than every 5 minutes). Denote by $\mathcal{J}^*(U)$ the optimal expected profit given that the policy is chosen from $\Pi(U)$. It is straightforward to verify that the following deterministic relaxation is an upper bound of $\mathcal{J}^*(U)$:

$$
\begin{aligned}
(\mathbf{DCP}(\zeta))(U)\quad \mathcal{J}(\zeta,U) := \max_{\boldsymbol{\lambda}}\quad & \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} r_{t,ij}(\lambda_{t,ij}) \\
s.t.\quad & \sum_{j=1}^{N}\sum_{s=1}^{t}\lambda_{s,ij} - \sum_{j=1}^{N}\sum_{s=1}^{(t-\tau_{ji})^+}\lambda_{s,ji} \le C_i, \forall i,t && \text{(EC.13a)}\\
& \zeta \le \lambda_{t,ij} \le 1-\zeta, \forall t,i,j && \text{(EC.13b)}\\
& \lambda_{s,ij} = \lambda_{t,ij}, \forall s,t\in\mathcal{U}_k, \forall k && \text{(EC.13c)}
\end{aligned}
$$

The dimension of the problem above can be reduced by eliminating the last constraint (i.e. the piecewise stationarity constraint). Define $\mathcal{U}^\dagger := \{1\le t\le T : t = k\cdot U \text{ for some } k\in\mathbf{N}_+\}$ and $\mathcal{U}_{ij} := \{1\le t\le T : t-\tau_{ij} = k\cdot U \text{ for some } k\in\{0\}\cup\mathbf{N}_+\}$ for each $i,j\in[N]$. For a fixed i, constraint (EC.13a) only needs to be considered on the set $\mathcal{U}_i = \mathcal{U}^\dagger\cup(\cup_{j=1}^N\mathcal{U}_{ji})$. To see why this is the case, consider a period $t\in[T]\setminus\mathcal{U}_i$. Let $\{u_k\}_{k=1}^K$ be the elements in $\mathcal{U}_i$ indexed from smallest to the largest. If $t<u_1$, it is not difficult to check that constraint (EC.13a) at t is implied by constraint (EC.13a) at u_1, since $u_1 = \min\{U, \min_{j=1}^N\tau_{ji}\}$. Suppose that $u_k<t<u_{k+1}$. Note that, the difference between the LHS of constraint (EC.13a) at t and constraint (EC.13a) at $t+1$ is simply the net outflow, given by

$$
\Delta_{it} := \sum_{j=1}^{N}\lambda_{t+1,ij} - \sum_{j:t\ge\tau_{ji}}\lambda_{t+1-\tau_{ji},ji}
$$

By definition of $\mathcal{U}_i$, it is not difficult to see that, for any $u_k<s<v\le u_{k+1}$, $\lambda_{s,ij} = \lambda_{v,ij}$ for any i and j. Therefore, we know that $\Delta_{i,u_k} = \Delta_{i,u_k+1} = \cdots = \Delta_{i,u_{k+1}-1}$. As a result, the LHS of constraint (EC.13a) at t is smaller than the LHS of constraint (EC.13a) at u_k if $\Delta_{it}<0$, and is smaller than the LHS of constraint (EC.13a) at u_{k+1} if $\Delta_{it}>0$. The argument above suggests that, after variable reduction, (EC.13a) can be equivalently reformulated as follows:

$$
\begin{aligned}
\max_{\boldsymbol{\lambda}}\quad & \sum_{k=1}^{T/K}\sum_{t\in\mathcal{U}_k}\sum_{i=1}^{N}\sum_{j=1}^{N} r_{t,ij}(\lambda_{k,ij}) \\
\text{s.t.}\quad & C_i \ge \sum_{j=1}^{N}\left(\sum_{k=1}^{\lfloor t/U\rfloor} U\cdot\lambda_{k,ij} + (t-U\cdot\lfloor t/U\rfloor)\,\lambda_{\lceil t/U\rceil,ij}\right) && \text{(EC.14a)}
\end{aligned}
$$

$$-\sum_{j:\tau_{ji}<t}\left(\sum_{k=1}^{\lfloor (t-\tau_{ji})^+/U\rfloor} U\cdot\lambda_{k,ji}+\left((t-\tau_{ji})^+-U\cdot\lfloor (t-\tau_{ji})^+/U\rfloor\right)\lambda_{\lceil (t-\tau_{ji})^+/U\rceil,ji}\right),\forall i,t\in\mathcal{U}_i$$

$$\zeta\leq\lambda_{k,ij}\leq 1-\zeta,\qquad \forall i,j,k \tag{EC.14b}$$

Note that, while the original optimization problem DCP has about NT constraints and N^2T decision variables, (EC.14) has at most $N^2(T/U)$ constraints and $N^2(T/U)$ decision variables. For the model calibrated using the Manhattan yellow taxi dataset, $T=7,200$ and $N=20$, which means the there are about $144,000$ constraints and $2,880,000$ decision variables and the original DCP is intractable using standard desktops. If we restrict the baseline optimal demand rate to be piecewise stationary for each 5-minute interval, (EC.14) has about $48,000$ constraints and $48,000$ decision variables. Using fmincon function in MATLAB, problem of such scale can be solved in hours.

EC.4. Performance Analysis of R-ABC: Proof of Theorem 4

Let $\pi=$R-ABC. Since the proof is similar to the proof of Theorem 3, for brevity, we will primarily focus on the differences. By an argument that is similar with the Proof of Lemma 2, we know that there exist some positive constants $\bar{\zeta}$ and Ψ_0 such that, for any $\zeta,\zeta^q<\bar{\zeta}$, we have $\mathcal{J}^R(0,0)-\mathcal{J}^R(\zeta,\zeta^q)\leq\Psi_0N^2T(\zeta+\zeta^q)$. We further define $\eta=(\epsilon\underline{\tau}^n-1)/8$ and $\eta^q=(\epsilon^q\underline{\tau}^n-1)/8$. We prove the results by considering two cases.

Analysis for small n

By definition, we know that both ζ and η/b converges to zero as n goes to infinity. Therefore, define $\Omega:=\max\{n\in\mathbb{Z}_{++}:(\underline{\tau}^n)^{2/3}\leq\max\{8N^2/(T^n)^2,512\log T^n,256\sqrt{\log T^n}/\bar{\zeta}\}\}$ (if the right hand side is an empty set, let $\Omega:=0$). Similar to the proof of Theorem 2, when $n\leq\Omega$, $\mathcal{L}^\pi(n)\leq M_1(\log T^n)^{2/3}/(\underline{\tau}^n)^{2/3}$ where M_1 is independent of n, T, and τ_{ij}.

Analysis for large n

When $n>\Omega$, by the definition of Ω, (17) and (18), the following holds:

$$\textbf{C3}:\quad \bar{\zeta}>\max\{\zeta,\zeta^q\},\zeta=2\epsilon,\zeta^q=2\epsilon^q,\text{ and }\frac{2\eta}{b}<1. \tag{EC.15}$$

Define the following sequence of events for all arc $i\to j$ and batch k

$$\mathcal{R}_{ijk}:=\left\{\max_{t\in\mathcal{T}_{ij}^k}\left|\sum_{s=\min\{v\in\mathcal{T}_{ij}^k\}}^{t}\delta_{s,ij}^n\right|<\eta\right\}\cap\left\{\max_{t\in\mathcal{T}_{ij}^{q,k}}\left|\sum_{s=\min\{v\in\mathcal{T}_{ij}^{q,k}\}}^{t}\delta_{s,ij}^{q,n}\right|<\eta^q\right\}$$

Further define $\mathcal{R}:=\cap_{i,j,k}\mathcal{R}_{ijk}$. We stat a lemma that says that condition **C3** implies that supply at all the nodes will not run out throughout the planning horizon with high probability. In fact, this lemma should be considered as the analogue of Lemma 4 in the analysis of ABC.

Lemma EC.2. *If **C3** holds, then for all sample paths on $\mathcal{R}$, the following condition holds for all t,*

$$\mathbb{H}_t: \quad C^n_{t,i} \geq N \text{ for all } i, \text{ and } p^\pi_{t,ij} = p^n_{t,ij}(\hat{\lambda}_{t,ij}) \text{ and } \mathbf{E}^\pi[Q^\pi_{t,ij}|\mathcal{F}_t] = \hat{q}_{t,ij} \text{ for all } i \to j.$$

where $\hat{\lambda}_{t,ij} = \lambda^{n,\zeta,\zeta^q,D}_{t,ij} - \epsilon - u_{t,ij} \cdot \bar{\delta}^{\kappa_{ij}(t)-1}_{ij} \in (0,1)$ *and* $\hat{q}_{t,ij} = q^{n,\zeta,\zeta^q,D}_{t,ij} - \epsilon^q - u^q_{t,ij} \cdot \bar{\delta}^{q,\kappa^q_{ij}(t)-1}_{ij} > 0$. *Moreover,* $\mathbf{P}(\mathcal{R}) \geq 1 - 4N^2 \cdot b^{-1}(T^n)^{-3}$.

Proof. First, it is straightforward to verify that, on all sample paths on $\mathcal{R}$, it holds that $\hat{\lambda}_{t,ij} \in (0,1)$ and $\hat{q}_{t,ij} > 0$. In particular, when $\bar{\delta}^{q,\kappa^q_{ij}(t)-1}_{ij} > 0$, $\hat{q}_{t,ij} > 0$ holds trivially; if $\bar{\delta}^{q,\kappa^q_{ij}(t)-1}_{ij} < 0$, we further have

$$\begin{aligned}
\hat{q}_{t,ij} &= \frac{q^{n,\zeta,\zeta^q,D}_{t,ij}}{2} - \epsilon^q + \frac{q^{n,\zeta,\zeta^q,D}_{t,ij}}{2} - \frac{q^{n,\zeta,\zeta^q,D}_{t,ij}}{\sum_{v\in\mathcal{T}^{q,k}_{ij}} q^{n,\zeta,\zeta^q,D}_{v,ij}} \cdot \bar{\delta}^{q,k-1}_{ij} \\
&\geq \frac{\zeta^q}{2} - \epsilon^q + q^{n,\zeta,\zeta^q,D}_{t,ij} \cdot \left(\frac{1}{2} - \frac{\eta^q}{b}\right) > 0,
\end{aligned}$$

where the last inequality holds by condition **C3**. We now proceed to prove condition $\mathbb{H}_t$ by induction. When $t = 1$, $C^n_{1,i} = nC_{1,i} \geq N$ by definition, which also implies the rest two identities by the definition of R-ABC. This completes the inductional basis. Suppose that $\mathbb{H}_s$ holds for all $s \leq t$, we now show that $\mathbb{H}_{t+1}$ holds as well. At the beginning of period $t+1$, we know that

$$C^n_{t+1,i} = C^n_{1,i} - \sum_{j=1}^{N}\sum_{s=1}^{t}(D^n_{s,ij} + Q^n_{s,ij}) + \sum_{j=1}^{N}\sum_{s=1}^{(t-\tau^n_{ji})^+}(D^n_{s,ji} + Q^n_{s,ji})$$

By a very similar derivation as in the proof of (EC.10) and (EC.11), it is straightforward to verify that $C^n_{t+1,i} \geq N\underline{T}^n(\epsilon + \epsilon^q) - 4N(\eta + \eta^q) = N$, which implies that $p^\pi_{t,ij} = p^n_{t,ij}(\hat{\lambda}_{t,ij})$ and $\mathbf{E}^\pi[Q^\pi_{t,ij}|\mathcal{F}_t] = \tilde{q}_{t,ij} = \hat{q}_{t,ij}$ by the definition of R-ABC. At last, we prove the probability bound. Similar to the proof of Lemma 4, we will apply the union bound to obtain an upper bound on the probability of the complement of $\mathcal{R}$. Since the upper bounds on the demand variability has been derived already, we only need to focus on the variability due to sampling the relocation quantity. For any batch $\mathcal{T}^{q,k}_{ij}$, we have

$$\begin{aligned}
&\mathbf{P}\left\{\max_{t\in\mathcal{T}^{q,k}_{ij}}\left|\sum_{s=\min\{v\in\mathcal{T}^{q,k}_{ij}\}}^{t}\delta^{n,q}_{s,ij}\right| \geq \eta^q\right\} \\
&\leq 2\exp\left(\sum_{s\in\mathcal{T}^{q,k}_{ij}}\min\left\{q^{n,\zeta,\zeta^q,D}_{t,ij} - \lfloor q^{n,\zeta,\zeta^q,D}_{t,ij}\rfloor, \lceil q^{n,\zeta,\zeta^q,D}_{t,ij}\rceil - q^{n,\zeta,\zeta^q,D}_{t,ij}\right\}\cdot r^2 - \eta^q\cdot r\right) \\
&\leq 2\exp\left(\left(\sum_{s\in\mathcal{T}^{q,k}_{ij}} q^{n,\zeta,\zeta^q,D}_{t,ij}\right)\cdot r^2 - \eta^q\cdot r\right) \\
&\leq 2\exp\left((b+1)\cdot r^2 - \eta^q\cdot r\right) = 2\exp\left(-\frac{(\epsilon^q\underline{T}^n - 1)^2}{256(b+1)}\right) = 2(T^n)^{-4}
\end{aligned}$$

where the first inequality follows from (EC.1), the third inequality holds by the definition of $\mathcal{T}_{ij}^{q,k}$, the first equality holds by setting $r = \eta^q/(2(b+1)) \in (0,1)$ (the inclusion follows by **C3**), the last equality follows by the definition of ϵ^q. Combining the bound above with the upper bounds on the demand variability gives us the probability upper bound. $\square$

We now move on to bound the profit loss. Define two positive constants $Q_{\max} := \max_i C_i/T$ and $c_{\max} := \max_{t,i,j} c_{t,ij}$. It is straightforward to check that the feasible relocation quantity in each period is upper bounded by $NQ_{\max}T^n$. Similar to (12), we have

$$\begin{aligned}
&\mathcal{J}^R(n,\zeta,\zeta^q) - \mathbf{E}^\pi\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} p_{t,ij}^\pi D_{t,ij}(p_{t,ij}^\pi) - c_{t,ij}Q_{t,ij}^\pi\right] \\
&\le \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[r_{t,ij}^n(\lambda_{t,ij}^{n,\zeta,D}) - r_{t,ij}^n(\hat{\lambda}_{t,ij})|\mathcal{R}\right] + r_{\max}\cdot N^2T^n\mathbf{P}(\mathcal{R}^c) \\
&\qquad \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[c_{t,ij}(\hat{q}_{t,ij} - q_{t,ij}^{n,\zeta,\zeta^q,D})|\mathcal{R}\right] + c_{\max}Q_{\max}N^3(T^n)^2\mathbf{P}(\mathcal{R}^c)
\end{aligned} \tag{EC.16}$$

The first and the second term can be bounded using the same argument as in the proof of Theorem 3. As for the second term, since the sampling procedure across periods are independent, we have $\mathbf{E}^\pi[\bar{\delta}_{ij}^{q,\kappa_{ij}^q(t)-1}] = 0$. Therefore, we further know that

$$\begin{aligned}
&\sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[c_{t,ij}(\hat{q}_{t,ij} - q_{t,ij}^{n,\zeta,\zeta^q,D})|\mathcal{R}\right] \le 0 - \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{E}^\pi\left[c_{t,ij}u_{t,ij}^q\cdot\bar{\delta}_{ij}^{q,\kappa_{ij}^q(t)-1}\middle|\mathcal{R}\right] \\
&= \sum_{t=1}^{T^n}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{P}(\mathcal{R})^{-1}\cdot\mathbf{E}^\pi\left[c_{t,ij}u_{t,ij}^q\cdot\bar{\delta}_{ij}^{q,\kappa_{ij}^q(t)-1}\middle|\mathcal{R}^c\right]\mathbf{P}(\mathcal{R}^c) \le \frac{4c_{\max}N^2}{b\zeta^qT^n(1-2N^2b^{-1}(T^n)^{-2})}
\end{aligned}$$

where the second inequality follows by Lemma EC.2 and the fact that $u_{t,ij}^q \le 1$, $|\delta_{t,ij}^q| < 1$ almost surely and $q_{t,ij}^{n,\zeta,\zeta^q,D} \ge \zeta^q$. Combining the bounds above with the upper bound on the revenue loss (see the proof of Theorem 3), we have

$$\begin{aligned}
\mathcal{L}^\pi(n) &= (\mathcal{J}^R(n,0,0) - \mathcal{J}^R(n,\zeta,\zeta^q) + \mathcal{J}^R(n,\zeta,\zeta^q) - \mathbf{E}^\pi[R^\pi(n)])/T^n \\
&\le \Psi_0(\zeta+\zeta^q) + N^2\left[\Psi_3\left(\epsilon + 2\epsilon^2 + \frac{64\log T^n}{b}\right) + \frac{4(\Psi_3 + c_{\max})N^2}{b\zeta T^n(1-2N^2b^{-1}(T^n)^{-2})} + \frac{2N^2(r_{\max} + c_{\max}Q_{\max}N)}{b}\right] \\
&\le M_2\cdot\frac{1+(\log T^n)^{\frac{3}{2}}}{(\underline{T}^n)^{2/3}}
\end{aligned}$$

where $M_2 = 64\Psi_0 + 160N^2\Psi_3 + 4N^4(\Psi_3 + c_{\max}) + 2N^4(r_{\max} + c_{\max}Q_{\max}N)$ is independent of n. Setting $\tilde{\Psi}_3 = \max\{M_1, M_2\}[(\log(2))^{-3/2} + (1+\log(T)/\log(2))^{3/2}]\underline{T}^{-2/3}$ completes the proof. $\square$