MICHIGAN ROSS

Working Paper

Real-Time Dynamic Pricing for Revenue Management with Reusable Resources and Deterministic Service Time Requirements

Yanzhe (Murray) Lei
Stephen M. Ross School of Business
University of Michigan

Stefanus Jasin
Stephen M. Ross School of Business
University of Michigan

Ross School of Business Working Paper Series
Working Paper No. 1327
July 2016

Real-Time Dynamic Pricing for Revenue Management with Reusable Resources and Deterministic Service Time Requirements

Yanzhe (Murray) Lei · Stefanus Jasin
{leiyz, sjasin}@umich.edu
Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI 48109

We consider the setting of a firm that sells a finite amount of resources to price-sensitive customers who arrive randomly over time according to a specified non-stationary rate. Each customer requires a service that consumes one unit of resource for a *deterministic* amount of time, and the resource is *reusable* in the sense that it can be immediately used to serve a new customer upon the completion of the previous service. The firm's objective is to set the price dynamically to maximize its expected total revenues. This is a fundamental problem faced by many firms in many industries. We formulate this as an optimal stochastic control problem and develop two heuristic controls based on the solution of the deterministic relaxation of the original stochastic problem. The first heuristic control is static since the corresponding price sequence is determined before the selling horizon starts; the second heuristic control is dynamic, it uses the first heuristic control as its baseline control and adaptively adjusts the price based on previous demand realizations. We show that both heuristic controls are asymptotically optimal in the regime with large demand and large number of resources. Finally, we consider two important generalizations of the basic model to the setting with multiple service types requiring different service times and the setting with advance service bookings.

Key words: dynamic pricing; reusable resources; asymptotic analysis

---

1. Introduction

Consider a firm managing a fixed amount of resources to satisfy time-varying price-dependent demand over a finite (selling) horizon. The resources are homogeneous, which means that customers do not have preference over a specific unit of resource, and each arriving customer requests a single unit of resource for a consecutive and *deterministic* amount of time (i.e., deterministic *service time*).

If a resource is available at the time of a new arrival, the new customer is immediately admitted into the system at the current list price and the service is immediately started without delay. (Later in this paper we will also consider the case with advance service booking where the service can be started at a fixed future time.) After the service is completed, the corresponding resource is released and can be directly used to satisfy a new demand (i.e., resource is *reusable*). The firm's objective is to maximize her expected total revenues throughout the horizon by setting prices dynamically. This is a fundamental problem faced by many firms in many different industries and the nature of this problem is not exactly identical to the canonical revenue management problem with stochastic demand and limited inventory (e.g., the classic model proposed in Gallego and Van Ryzin 1997). (To be precise, although it is mathematically possible to model revenue management with reusable resources and deterministic service time requirement using the same modeling approach as in the classic revenue management literature, the scale of the problem primitives for the applications considered in this paper is different from that considered in the standard revenue management literature. Hence, a different approach is needed to properly analyze this model; see Section 3 for more discussions.) Our main contribution in this paper is in developing a real-time dynamic pricing control that is easy to implement and has a provably good performance. We first show how to do this for a basic setting with one service type and immediate service requirement; we then show how our idea can be applied to more complicated settings with multiple service types with heterogenous service time requirements and advance service booking. We believe that the idea behind our proposed control can be potentially used to develop more sophisticated dynamic pricing controls for other complicated real-world problems.

Our formulation captures the critical operational trade-offs faced by firms in different industries. On the one hand, capacity needs to be sufficiently utilized throughout the selling horizon since, at any point of time, any unused or idle capacity constitutes immediate monetary loss; on the other hand, firms may also want to ration the capacity to anticipate potential peak periods in the future where the system is fused with incoming demands. The main challenge here is how to properly balance the capacity utilization during different *service cycles*. (The meaning of service cycle will be explained in Section 3. Note that, due to the difference in the scale of problem primitives as noted above, the classic revenue management problem effectively only has one service cycle as demands are typically modeled to be fulfilled only at the end of the selling horizon instead of on a rolling horizon basis. This is in contrast to the setting considered in this paper, which may have a large number of service cycles.) Although different cycles may appear to be independent of each other, they are connected through the realization of capacity utilization since capacity is finite and the utilization in one cycle affects the utilization in the subsequent cycle. This calls for a carefully designed dynamic pricing control to properly manage capacity utilization across different cycles.

In the queueing literature, a finite capacitated system similar to the one considered in our work is often termed as a *loss system*, since an arriving customer is rejected when the capacity is full (on the contrary, in a *delay system* model, incoming customers are allowed to wait in a queue, see Hampshire et al. 2009). Many firms providing virtual services such as telecommunication, smart grid, and Internet-based service (Voice-over-IP, wireless data transfer) can be appropriately modeled as loss systems. In all these examples, pricing decision is important not only because it serves as a marketing instrument that determines the total revenues collected by the firm, but also as a control instrument by which the firm continuously manages the utilization level of her finite resources. There are at least two salient features of the firms' operation problem that often complicate the pricing decision: on the demand size, demand rates tend to change dynamically and is better described as a time-inhomogeneous process (see Brown et al. 2005 for a statistical study in the setting of call center); on the supply side, capacity expansion is sometimes a long-term investment decision and the current capacity is not easily scalable within a short period of time, thus, they must be managed properly. The effectiveness of dynamic pricing in matching time-varying demand with limited capacity has been widely recognized and implemented by many firms, e.g., mobile service providers in Africa charges rates of voice-call dynamically to alleviate the burden on their bandwidth during peak periods and stimulate demand during low period

(Economists 2009); smart grids in United States and Europe experiment with programs that bill customers' consumption of electricity on a time-dependent rate (Hu et al. 2015).

In addition to the two salient features mentioned above, pricing decision is also often complicated by the fact that, once used, the same resource may continue to be used during a fixed period of time, and different customers may request to use the resource for different length of time (i.e., different service time). One of the emerging business that fits this feature is cloud computing, where firms deliver on-demand internet-based computing service to customers. Cloud computing service providers usually have a fixed amount of computation resources and lease their available resources to customers who arrive (either on spot or under subscription) randomly with specific request on usage time and capacity requirements. In the provision of cloud service, researchers and practitioners have advocated the economic benefit of dynamic pricing strategy for many cloud service settings. By and large, dynamic pricing has been implemented under the form of utilization-based pricing (CloudSigma, Jelastic, PiCloud), real-time bidding (Amazon Elastic Computing Cloud (EC2) Spot), and many others (Al-Roomi et al. 2013). As arguably the largest cloud computing service provider, Amazon launched its EC2 Spot service in 2009, whose per-hourly price is determined in real-time by a Vickrey-style auction. More specifically, after customers submit sealed bids, Amazon will computes a market clearing price (a.k.a "spot price"). All customers with bid above spot price win, and pay the lowest winning bid for the service with the requested features such as duration, memory size, etc. Not too surprisingly, the resulting price trajectory is often highly non-stationary (Xu and Li 2013) and, in spite of its flexibility, the implementation of bidding mechanism has its own flaws. As an example, Cheng et al. (2016) shows empirically that, for the same type of computing service on Amazon EC2 Spot platform, network latency causes significant and consistent price difference between its East and West data center, which clearly opens an arbitrage opportunity. These problems would not have existed if the firm has a full control over the price trajectory. The key technical question is how to implement a dynamic pricing in a way that matches time-varying demand with fixed but reusable resources. This has motivated many researchers to investigate a proper dynamic pricing control under various settings where service providers fully control the price (e.g., Xu and Li 2013, Alzhouri and Agarwal 2015, Arshad et al. 2015).

Other than the examples discussed above, many firms managing physical resources are also well described by our model, including some classic examples that are well-known for their adoptions of dynamic pricing such as car rental and hotel reservation. (Due to the reusable nature of their resources, both car rental and hotel reservation are more properly modeled using the framework of revenue management with reusable resources and deterministic service time requirements instead of using the classic revenue management framework motivated by airline application.) Moreover, our model can also be used to address the demand-supply matching problem faced by many emerging so-called on-demand service firms. These firms usually control a finite number of resources and offer them to be consumed by customers who book services (either in advance or on the spot) through internet or smartphone. Industries that have seen the booming of on-demand service providers include vehicle rental (Zipcar, Citi-Bike), logistic (Project44), food delivery (Instacar, Sprig), car parking (Luxe), and beauty service (StyleSeat) (Bensinger 2015). One prevalent feature in many on-demand service firms is that demand is characterized by both a specified service type and an intended usage time, including the starting and ending times of the service. Moreover, since customers mostly interact with the firm using digital platforms, existing digital user interfaces often enable the firm to effortlessly manage demand by dynamically changing prices. Indeed, some firms have already used dynamic pricing on a daily basis. For example, Project44 provides dynamic pricing solutions to third-party logistics company owning their own trucks and facilities (Project44 2015); Tock provides ticketing systems to high-end restaurants where reservation of seats are dynamically priced (Businessweek 2015); Sprig uses its own employees to deliver fresh made meals to customers at a delivery fee that changes dynamically (Chamlee 2016). Other firms that have not yet deployed dynamic pricing have also acknowledged its advantage: According to Robin Chase, the founder of Zipcar, utilizing data to correctly and dynamically set the price on car-sharing platform can largely increase the efficiency and sustainability of the deployment of city services (GreenBiz 2014). Thus, although not every firms under the banner of on-demand economy is currently using dynamic pricing, given its simplicity and good performance, we believe that our proposed real-time dynamic pricing control can provide a useful guidance on how to do real-time dynamic pricing when the firm finally needs it.

We want to re-emphasize that eventually different business models may have different complexities that require separate customized dynamic pricing solutions. In this paper, we simply focus on a simple model that captures the most fundamental aspects of dynamic pricing with reusable resources and deterministic service time requirements. We hope that our result can be used to design more sophisticated algorithms to be used in all of the aforementioned examples.

Our results and contributions. In this paper, we consider a multi-period dynamic pricing problem faced by a revenue-maximizing firm with finite reusable resources and deterministic service time requirements. Our analysis and results are summarized below:

1. We first consider a basic model where all customers have the same deterministic service time requirements, there is no delay in service fulfillment, and demand rate as a function of time and price is non-stationary. We propose a deterministic relaxation of the optimal control formulation, and show that its objective value serves as an upper bound for the optimal expected total revenues under the original stochastic control problem. This allows us to evaluate the performance of any feasible pricing control by its *average regret*, defined as the average (over T periods) difference between the optimal value of the deterministic formulation and the expected total revenues collected under the prescribed control.
2. Our first heuristic control, which we call *Deterministic Price Control (DPC)*, applies price p_t in period t in such a way that that the expected demand in period t equals the computed deterministic demand rate under the deterministic formulation minus a constant. The constant serves as a buffer on random error, for the purpose of hedging against uncertainty. The size of this buffer needs to be carefully chosen: It needs to be large enough such that the resource is not depleted too often; yet, it cannot be too large otherwise the total revenues collected by

the firm will deviate too far from the optimal one. We obtain a general bound on the average regret of DPC under arbitrary problem parameters and show that, under an optimal choice of buffer size, the average regret of DPC converges to zero at a rate of $\tilde{O}(n^{-\frac{1}{2}})$, where n is the size of the problem (i.e., the size of potential demand during a service cycle, which is to be defined later, and capacity are both of order n).

3. One drawback of DPC is that the price p_t to be applied during period t is already determined at the beginning of the selling horizon and it does not take into account to the realized demand observations during periods 1 to $t-1$. This suggests a room of improvement and motivates our second heuristic control, which we call *Deterministic Price Control with Batch Adjustment (DPC-Batch)*. DPC-Batch divides the selling horizon into batches of the same size. At each period, in addition to making sure that we have the buffer as in the case of DPC, we also set the price in such a way that the cumulatively demand errors (i.e., from expected demands) during the previous batch is uniformly corrected by the new demands in the current batch. We obtain a general bound for the average regret of DPC-Batch under arbitrary problem parameters and show that, under an optimal choice of buffer size and batch size, the average regret of DPC-Batch is of order $\tilde{O}(n^{-\frac{2}{3}})$, which significantly improves the performance of DPC. We conduct several numerical experiments that validate our theoretical findings.
4. Finally, we consider two extensions of the basic model to include two important features often found in practice, namely heterogeneous service time requirements (where different service type may require different service time) and advance service booking (where different service type may be started at different time in the future). We focus our analysis on the generalization of DPC-Batch. For the sake of clarity how the analysis of our basic model can be extended to a more general model, we treat these two extensions as separate instances instead of one. Under properly chosen problem parameters, we show that the average regret of the generalized DPC-Batch for each of these extensions is still of the order $\tilde{O}(n^{-\frac{2}{3}})$.

Organization of the paper. The related literature is reviewed in Section 2. In Section 3, we formulate the basic model of dynamic pricing with reusable resource, and discuss our performance measure. We propose and analyze a static heuristic control (DPC) and its dynamic improvement (DPC-Batch) in Sections 4 and 5, respectively. The performance of both DPC and DPC-Batch are tested in simple numerical experiments in Section 6. Sections 7 and 8 discuss two extensions of the basic model that allow heterogeneous service time requirements and advance booking. Finally, in Section 9, we conclude the paper. The proof of some of the results and the details of the numerical experiments can be found in the electronic companion for this paper.

2. Literature Review

Broadly speaking, our work is related to the extensive literature on dynamic pricing and revenue management, queueing and service operations, and on-demand service platforms. In terms of methodology, our work is related to the study of asymptotic performance of heuristic controls with real-time adjustment. We discuss them in turn.

Dynamic pricing and revenue management. Given the space limit, we will not attempt to discuss all the related literature but only highlight the most relevant works (interested readers are referred to the extensive surveys by Bitran and Caldentey 2003, Talluri and van Ryzin 2006 and Özer and Phillips 2012.) Instead, we discuss in details two papers that are most closely related to our work, both are motivated by the revenue management problem in cloud computing setting. Xu and Li (2013) study the dynampic pricing problem of a cloud service provider that leases resources to customers with exponential service time and price-dependent Poisson arrival. They obtain some structural properties for the capacitated system under stationary demand and also for the uncapacitated system under non-stationary demand. However, no dynamic pricing heuristic control is proposed and the optimal price is still time-consuming to compute, especially when demand is non-stationary. Our work complements their work: We explicitly address the capacitated system with non-stationary demand and deterministic service time requirement, and focus on developing an easy-to-implement heuristic control instead of studying the properties of the optimal solution. Borgs et al. (2014) study a similar problem under non-stationary demand with limited time-varying capacity and customers' strategic waiting. In their model, demands are assumed to be deterministic and the price trajectory for the whole season is announced at the beginning of the horizon. They show that the resulting optimization problem is non-convex and propose a dynamic programming-based algorithm that can be run in polynomial time. The key difference between our model and theirs is on the stochasticity of demand and customer's strategic waiting: In our model, demand is random and, thus, an adaptive heuristic control is needed to guarantee a near-optimal revenue. (In many service settings, especially for the on-demand platform, uncertainty in demand pervasively exists and introduces a significant difficulty in control design.) Unlike their model, we do not explicitly consider customer waiting behavior in our current work. Although customers' waiting is an important issue and needs to be properly taken into account when designing a dynamic pricing control, proposing a provably good heuristic control under a combination of stochastic demand, limited inventory, and customers' waiting is a notoriously difficult problem even in the traditional revenue management setting (see e.g., Liu and Cooper 2015 and Chen and Farias 2016 for recent progress). Thus, we leave this for future research pursuit.

Queueing and service operations. As explained in the previous section, our model is similar to the loss system in the queueing literature. Pricing decision in such model has been studied extensively under various setting (e.g., Lanning et al. 1999, Courcoubetis et al. 2001 and Maglaras and Zeevi 2005). Most of these papers propose heuristic controls based on a fluid approximation of the original stochastic control problem under the assumption of stationary arrival and exponential service time. An exception to this is Hampshire et al. (2009), where demand follows a non-homogeneous Poisson process and the firm has to satisfy a Quality-of-Service constraints which requires the blocking probability to be bounded. They develop a dynamic pricing control using deterministic optimal control theory and show numerically that this control performs better than static or myopic pricing control; however, no theoretical performance guarantee is provided of their proposed control. Another major stream of literature studies the property of the optimal admission control of loss system, including Miller (1969), Kelly (1991), Altman et al. (2001), Örmeci et al. (2001), Savin et al. (2005), Gans and Savin (2007), Papier and Thonemann (2010) and Jain et al.

(2015). Yet, none of them consider the design of practical and provably-good heuristic controls. There are two exceptions: Levi and Radovanovic (2010) propose a heuristic control based on a knapsack-type linear program and show the asymptotic optimality their proposed control under a general service time distribution, and Levi and Shi (2015) generalize this heuristic control to the setting with advance booking and provide an asymptotic upper bound on the blocking probability. However, both Levi and Radovanovic (2010) and Levi and Shi (2015) assume stationary demand and do not consider dynamic pricing.

Aside from the literature on loss model, dynamic pricing has also been studied in the literature on delay model. From the modeling perspective, researchers that study optimal dynamic pricing control either assume that customers are sensitive to price only but not delay (e.g., Low 1974, Paschalidis and Tsitsiklis 2000, Yoon and Lewis 2004, Maglaras 2006) or customers are sensitive to both price and delay (e.g., Chen and Frank 2001, and Ata and Shneorson 2006, Afèche and Ata 2013). Several papers study asymptotically optimal dynamic pricing controls: Çelik and Maglaras (2008) and Ata and Olsen (2009, 2013) study a revenue maximizing control when the firm dynamically quotes lead-times; Besbes and Maglaras (2009) study dynamic pricing where the market size varies stochastically over time; assuming observable queue length and stochastic customer valuation, Kim and Randhawa (2015) propose a heuristic control that continuously refines the baseline control given by a fluid approximation, and show (somewhat surprisingly) that the average regret is on the order of $\tilde{O}(n^{-\frac{2}{3}})$. (To the best of our knowledge, Kim and Randhawa (2015) is the only work in the queueing literature that shows dynamic pricing can achieve an average regret with order smaller than the more typical $\tilde{O}(n^{-1/2})$.) Aside from not permitting customers to wait, our model is different from the above cited works adopting asymptotic analysis in two aspects: (1) We assume that the service time is deterministic (earlier works assume that it is exponentially distributed) and the demand function can vary over time (earlier works assume a stationary willingness-to-pay distribution), and (2) we also consider an extension with advance service booking. The appropriateness of using either a deterministic service time or exponentially distributed service time is dictated by the application context. In this paper, we choose to work with deterministic service time because, in most of the applications that we are considering, service process is not memoryless (as would have been implied by an exponentially distributed service time). Thus, our work complements existing works in the queueing literature by developing a near-optimal dynamic pricing control that can be applied in the setting of non-stationary demand, deterministic service time requirements, and advance service booking. Moreover, we also complement the result of Kim and Randhawa (2015) by showing that the $\tilde{O}(n^{-2/3})$ bound is also achievable in our setting.

On-demand service platform. Our paper is also connected to the growing literature on the operational problems faced by firms providing various types of on-demand services. Most of the existing works focus on a specific industry and, henceforth, deal with more complicated models than ours. (Per our discussions in Section 1, our objective in this paper is to focus on the most fundamental aspects of revenue management with reusable resources and deterministic service time requirements instead of addressing a particular problem instance with all its complexities.) One line of research in this literature studies the logistic optimization problems for vehicle/bike sharing platforms (e.g., Raviv and Kolka 2013, Shu et al. 2013, Schuijbroek et al. 2013, O'Mahony 2015 and Kaspi et al. 2016). Existing works that study pricing decisions are Pfrommer et al. (2014) and Waserhole (2014). They both consider a network of shared mobility system and view price as an incentive to direct customers to allocate resources in a way that inventory balancing is properly maintained throughout the network. Different heuristic controls are proposed based on certainty equivalent principle and are tested using numerical experiments. In contrast to our work, Pfrommer et al. (2014) and Waserhole (2014) use platform's expected cost of repositioning vehicle as the objective. Another stream of literature studies the optimization of dynamic delivery fee for the attended home delivery firms, e.g., Campbell and Savelsbergh (2006), Asdemir et al. (2009), Klein et al. (2015). The key trade-off addressed in these works is how to use price to incentivize customers to allocate their demands to different delivery time slots such that the profit (delivery fee minus the cost associated with service type and time slots) is maximized. Moreover, their systems are capacitated in the sense that the delivery capacity within each time slots is fixed and known. In comparison to our model, this modeling framework embraces less uncertainty since, in our model, the available capacity at any time depends dynamically on the past demand realizations.

Real-time control. In the broader dynamic optimization literature where a multi-period stochastic control problem is often difficult (if not impossible) to solve optimally, researchers often resort to simple heuristic controls. A specific type of heuristic control, called real-time control, calculates the decision at the current period as a simple (e.g., affine) function of a baseline control and the historical information. Driven by its practicality (as the name suggests, a real-time control adaptively adjusts the control on the fly and does not require heavy re-optimizations) and good performance, real-time control has been investigated in various fields, including robust optimization (Ben-Tal et al. 2004, Bertsimas et al. 2010), portfolio management (Calafiore 2009, Moallemi and Saglam 2012), and revenue management (Atar and Reiman 2012, Golrezaei et al. 2014, Chen et al. 2014, Lei et al. 2016). Closest to our paper are Jasin (2014) and Chen et al. (2015). They both consider the discrete-time version of the canonical dynamic pricing problem studied in Gallego and Van Ryzin (1997), and propose real-time price controls with provable performance guarantees. As discussed in Section 3.2, in theory, our problem can also be formulated using the same framework as in Gallego and Van Ryzin (1997); however, the reusability of resource in our setting introduces a non-trivial subtlety that prohibits a simple adoption of the heuristic controls proposed in Jasin (2014) and Chen et al. (2015) into our setting. (In fact, we show numerically in Section 6 that a simple adoption of this heuristic control performs very poorly.) Thus, our work complements existing works in the literature of real-time control by proposing a different real-time price control that is appropriate for the setting of revenue management with reusable resources and deterministic service time requirements.

3. Basic Model

In this section, we first discuss the setting and primitive of our basic model. Next, we discuss the stochastic and deterministic formulations of our dynamic pricing problem. Finally, we discuss our performance measure.

3.1. The Setting

We consider a discrete-time model with T periods and C units of resource. (Although we assume a discrete-time model, our results also hold for a continuous-time model with Poisson arrivals.) For our basic model, we assume that the firm only sells one service (or product) type where each request requires one unit of resource and n units of service time (or n periods). For example, if $n = 1$, then the service started in period 1 is completed at the end of period 1 and the resource used to fulfill this service is immediately available to fulfill a new request in period 2. Demand rate, as a function of price, in period t is given by $\lambda_t(p_t)$, and the corresponding revenue rate is given by $r_t(p_t) = p_t \cdot \lambda_t(p_t)$. Let $D_t(p_t)$ denote the realized demand in period t under price p_t. By definition, we have $\mathbf{E}[D_t(p_t)] = \lambda_t(p_t)$ and $\mathbf{E}[p_t \cdot D_t(p_t)] = r_t(p_t)$. It is typically assumed in the literature that demand rate is invertible in price (see Assumption A1 below). Thus, by abuse of notation, we will also write $D_t(\lambda_t) = D_t(p_t(\lambda_t))$ and $r_t(p_t) = p_t \cdot \lambda_t(p_t) = \lambda_t \cdot p_t(\lambda_t) = r_t(\lambda_t)$ to denote the direct dependency of realized demand and revenue rate on demand rate instead of on price (we use $p_t(\cdot)$ to denote the inverse of $\lambda_t(\cdot)$). We assume that demands across different periods are independent, but demand rate as a function of time may be non-stationary. As is typical in the revenue management literature (see e.g. Jasin 2014), we further assume that at most one request arrives during each period. (Thus, $\lambda_t(p_t)$ can be interpreted as the arrival probability of a new request in period t under price p_t.) This is without loss of generality since our analysis can also be applied to the setting where multiple requests arrive in each period. Let Ω_p and Ω_λ denote the convex feasible set of price and demand rate, respectively. (For simplicity, we assume the same feasible sets in all periods.) Below, we state some standard regularity conditions on $\lambda_t(\cdot)$ and $r_t(\cdot)$:

A1. $\lambda_t(p_t) : \Omega_p \to \Omega_\lambda$ is bounded, twice differentiable, and invertible.

A2. There exists a "turn-off" price $\bar{p}$ such that $p_t^k \to \bar{p}$ implies $\lambda_t(p_t^k) \to 0$ for all t.

A3. For all t, $\lambda_t^k \to 0$ implies $\lambda_t^k \cdot p_t(\lambda_t^k) \to 0$ for all feasible sequences $\{\lambda_t^k\}_{k=1}^{\infty}$.

A4. $r_t(\lambda_t)$ is bounded, strictly concave, and has a finite maximizer $\lambda_t^* \in \Omega_\lambda$.

The above assumptions are sufficiently general and are immediately satisfied by most commonly demand functions including linear, exponential, power, and logit. The existence of a turn-off price $\bar{p}$ allows the firm to effectively turn off demand whenever needed (e.g., when no resource is currently available). It should be noted that although the theoretical turn-off price can be infinite (e.g., for exponential demand function with $\lambda_t(p_t) = a \cdot e^{-p_t}$), since real-world price is never infinite, we can assume without loss of generality that $\bar{p} < \infty$. (To be precise, we can pick a sufficiently large $\bar{p}$ such that both $\lambda_t(\bar{p})$ and $r_t(\bar{p})$ are very small. The exact value of $\bar{p}$ does not affect our analysis.)

3.2. The Stochastic and Deterministic Formulations of Dynamic Pricing Problem

The dynamics of our pricing problem are as follows. First, a new request arrives at the beginning of period t with probability $\lambda_t(p_t)$. If a unit of resource is available, the service is immediately started (i.e., no waiting is allowed) and, once a service is started in period t, it will be completed at the end of period $t+n-1$. The corresponding resource is then immediately available for a new service in period $t+n$. No intervention or cancellation is allowed, i.e., neither the firm nor the customer can stop the service before it is completed. Since we assume at most one request arrives in each period, at most one service is completed at the end of any period.

Let Π denote the set of all non-anticipating controls (i.e, the control that decides the price at the beginning of period t using only the accumulated information up to, and including, the end of period $t-1$), and let p_t^π denote the price to be applied during period t under policy $\pi \in \Pi$. The optimal stochastic control formulation of our dynamic pricing problem is given below:

$$\mathbf{OPT}: \quad J^* = \left\{ \max_{\pi \in \Pi} \mathbf{E}\left[\sum_{t=1}^{T} r_t(p_t^\pi) \right] : \sum_{s=\max\{1,\, t-n+1\}}^{t} D_s(p_s^\pi) \le C \;\; \text{for all } t \le T \right\}$$

where the constraints must hold almost surely, or with probability one. To understand the intuition behind the above constraints, note that the number of units of resource available at the beginning of period t is given by $C - \sum_{s=\max\{1,\, t-n+1\}}^{t-1} D_s(p_s^\pi)$. Here, we only need to consider total demands in the previous $n-1$ periods because any resource being used in period $s < \max\{1,\, t-n+1\}$ must already complete its assigned service and is either at an idle state at the beginning of period t or currently being used to satisfy a new request arriving in period $s \in [\max\{1, t-n+1\}, t-1]$, where by abuse of notation we use $[t_1, t_2]$ to denote $\{t_1, t_1+1, \ldots, t_2\}$. For a new service to be started in period t, we must satisfy capacity constraint $D_t(p_t^\pi) \le C - \sum_{s=\max\{1, t-n+1\}}^{t-1} D_s(p_s^\pi)$, or equivalently $\sum_{s=\max\{1, t-n+1\}}^{t} D_s(p_s^\pi) \le C$. This explains our constraints in **OPT**.

The exact stochastic formulation **OPT** is in general difficult to solve due to the famous "curse of dimensionality" of Dynamic Programming (DP). Our focus in this paper is on the construction of near-optimal heuristic controls using the solution of a deterministic analogue of **OPT**. We define a deterministic optimization **DET** as follows:

$$\mathbf{DET}: \quad J^D = \left\{ \max_{p_t \in \Omega_p} \sum_{t=1}^{T} r_t(p_t) : \sum_{s=\max\{1, t-n+1\}}^{t} \lambda_s(p_s) \le C \ \text{ for all } t \le T \right\}.$$

The above formulation is sometimes called a *fluid* model in the literature (e.g., Atar and Reiman 2012). Since demand is invertible in price (by Assumption A1), we can also re-write **DET** using demand rates as the immediate decision variables instead of prices as follows:

$$\mathbf{DET}: \quad J^D = \left\{ \max_{\lambda_t \in \Omega_\lambda} \sum_{t=1}^{T} r_t(\lambda_t) : \sum_{s=\max\{1, t-n+1\}}^{t} \lambda_s \le C \ \text{ for all } t \le T \right\}.$$

One of the benefit of the above re-formulation is that the constraints are now linear in the decision variables and the objective is strongly concave by Assumption A4; so, **DET** can be efficiently solved using an off-the-shelf convex optimization solver. Note that the constraints in **DET** can be more compactly written as $A\lambda \le \mathbf{e} \cdot C$, where λ is a column vector of demand rates, $\mathbf{e}$ is a column vector of ones with an appropriate length, and A is an appropriate constant matrix. Although this compact formulation is similar to the canonical deterministic formulation in the standard revenue management literature (e.g., Gallego and Van Ryzin 1997), it is important to note that the size of matrix A in our setting scales with T whereas the size of matrix A in the standard literature is independent of T. This seemingly minor difference has an important, non-trivial, consequence in heuristic design. This is the reason why a different approach is needed to properly analyze the general revenue management with reusable resources and deterministic service time requirements.

Let $p^D := (p_t^D)_{t=1}^T$ denote the optimal solution of **DET**, and let $\lambda^D := (\lambda_t^D)_{t=1}^T$ denote the corresponding optimal demand rates (i.e., $\lambda_t^D = \lambda_t(p_t^D)$ for all t). Unlike in the standard revenue management setting where the optimal deterministic price is static (i.e., $p_t^D = p_1^D$ for all t) when demand rates are stationary (see e.g. Gallego and Van Ryzin 1997), the optimal solution of **DET** is not necessarily static even when demand rates are stationary (except for a special case T is a constant multiplicand of n). Below, we state additional assumptions on λ^D and the derivatives of revenue rate and price as functions of demand rate. There exist positive constants φ_L , φ_U, and Ψ such that:

A5. $[\lambda_t^D - \varphi_L, \lambda_t^D + \varphi_U] \subseteq \Omega_\lambda$ for all t.

A6. $|r_t'(\lambda)|$, $|r_t''(\lambda)|$, and $|p_t'(\lambda)|$ are bounded by Ψ on $[\lambda_t^D - \varphi_L, \lambda_t^D + \varphi_U]$ for all t.

The above assumptions are sufficiently general. Assumption A5 corresponds to the case where, at least in a deterministic world, the prices in all periods are neither too low that they collectively induce too many demands nor too high that they collectively induce too few demands. (This reflects what we find in most real-world settings as typical prices are neither extremely low nor outrageously high.) On another note, this assumption is also easily satisfied when λ_t^* lies in an interior of Ω_λ for all t, which is not at all uncommon given the strong concavity of $r_t(\cdot)$ as a function of λ_t. The boundedness of the derivatives of the revenue and price functions in an interior of Ω_λ as stated in Assumption A6 are also quite natural and easily satisfied by many demand functions including linear, exponential, power, and logit. Note that we only require that these derivatives are bounded in a certain compact subset of Ω_λ instead of the whole Ω_λ. The later is too restrictive and is not possible even for the case of power demand function $\lambda_t(p_t) = a \cdot p_t^{-b}$ since $r_t'(\lambda_t) \to \infty$ as $\lambda_t \to 0$.

The following lemma tells us that J^D is an upper bound of J^*. This result is analogous to a standard result in the revenue management literature (e.g., Gallego and Van Ryzin 1997), and its proof utilizes a simple argument using Jensen's inequality. We state it here for the sake of completeness.

Lemma 1. $J^* \leq J^D$.

One of the benefit of Lemma 1 is that it allows us to use J^D as a proxy for J^*. This is particularly useful for the purpose of evaluating the performance of different heuristic controls since J^* is not practically computable. We discuss this next.

3.3. Performance Measure and Asymptotic Regime

Let R^π denote the total revenues collected under policy π throughout T periods. We are interested in measuring the average expected total losses, or average regret, of a given control with respect to the optimal control. However, since the optimal control is not computable as mentioned above, we will use the deterministic upper bound as a proxy. We thus defined the average regret of a non-anticipating control $\pi \in \Pi$ as follows:

$$\text{AvReg}(\pi) = \frac{J^D - \mathbf{E}[R^\pi]}{T}.$$

Intuitively, since the expected total revenues throughout T periods under the optimal control scales linearly with T, the above definition of average regret captures the order of relative regret with respect to the optimal control. In this paper, we are particularly interested in the case where n is large and $C = \Theta(n)$. This can be interpreted as the setting where total potential demands during a service cycle is large and we have just enough resources to satisfy the demands in one cycle. (For completeness, in Remarks 1 and 3 in Sections 4 and 5, we also discuss what happens when $C = o(n)$; this can be interpreted as the setting where either resources are very scarce or the length of service time is very long. The remaining case where we have a lot more resources than what we need to satisfy demands in one service cycle, i.e., $n = O(C)$, is less interesting as it reduces our dynamic pricing problem into an unconstrained problem and we can simply apply $p_t = p_t(\lambda_t^*)$ for all t.) This is not uncommon and is motivated by many practical applications discussed in Section 1. As the size of n can be very large (i.e., at least hundreds or thousands), we focus in constructing heuristic controls that are near-optimal in the so-called *asymptotic regime*. We would like to note that the setting where n is large and $C = \Theta(n)$ is also similar to the standard asymptotic setting in the queueing literature (e.g., Kim and Randhawa 2015) where both the demand and service rates are scaled by the same large constant.

We say that a control $\pi \in \Pi$ is *asymptotically optimal* if $\frac{J^D - \mathbf{E}[R^\pi]}{T} \to 0$ as $n \to \infty$ for a suitable value of T, which may also scale with n. In this paper, we prove that both DPC and DPC-Batch are asymptotically optimal. However, as n increases, the average regret of DPC-Batch converges to 0 faster than the average regret of DPC. (For our basic model, the convergence rate of DPC-Batch is approximately $n^{-2/3}$ whereas the convergence rate of DPC is approximately $n^{-1/2}$.) For ease of exposition, throughout the remaining of the paper, we will always assume that $\frac{T}{n} \in \mathcal{Z}^+$.

4. Deterministic Price Control

In this section, we first introduce a simple heuristic control called *Deterministic Price Control* (DPC) and then we analyze its performance.

4.1. Control Description and Statement of Result

Let C_t denote the number of units of resource available at the beginning of period t before the firm sets a new price p_t. The formal definition of DPC is given below.

Deterministic Price Control with Parameter ϵ (DPC(ϵ))

Step 1. Solve **DET** and get λ^D.

Step 2. At the beginning of each t, do:

a. If $C_t \geq 1$, set $p_t = \hat{p}_t^D$ where

$$\lambda_t(\hat{p}_t^D) = \lambda_t^D - \frac{\epsilon}{n};$$

b. Otherwise, set $p_t = \bar{p}$.

Note that DPC is parameterized by $\epsilon > 0$, and ϵ needs to be chosen such that $\lambda_t^D - \frac{\epsilon}{n} \in \Omega_\lambda$ (otherwise, the second step in DPC(ϵ) is not well-defined). Since the targeted demand rate in period t under DPC(ϵ) is $\lambda_t = \lambda_t^D - \frac{\epsilon}{n}$, the total targeted average demands in n consecutive periods (i.e., one service cycle) is at most $C - \epsilon$, which means that we are essentially holding back ϵ units of resource. We do this for the purpose of hedging against uncertainty: If total realized demands in the previous n periods turn out to be higher than expected, then we still have an extra ϵ units of resource that can be immediately used to satisfy demand. (From a theoretical perspective, having a positive ϵ is useful in making the analysis of DPC more tractable, though it may not be necessary for the actual implementation. In Section 6, we numerically test what happens when we set $\epsilon = 0$.) The following theorem tells us the performance of DPC.

THEOREM 1. *There exists a constant $M_1 > 0$ such that for all T, C, n, and $\epsilon \in [1, n\varphi_L]$,*

$$\textsc{AvReg}(DPC) \leq M_1 \cdot \left[\frac{\epsilon}{n} + \frac{T}{n} \cdot \exp\left\{-\frac{(\epsilon-1)^2}{36\min\{C-\epsilon,\, n\}}\right\}\right]. \tag{1}$$

In particular, if $C = a \cdot n$ for some $a > 0$, then using $\epsilon = 1 + 6\sqrt{b \cdot n \cdot \log n}$ for some $b > 0$ yields

$$\textsc{AvReg}(DPC) = O\left(\sqrt{\frac{b \cdot \log n}{n}} + \frac{T}{n^{1+\frac{b}{\max\{1,a\}}}}\right). \tag{2}$$

The first bound in Theorem 1 is very general; it highlights the impact of T, n, C, and $\epsilon \in [1, n\varphi_L]$ on performance. As for the second bound, as long as T grows at a polynomial rate in n (i.e., T can be very large, especially when n is large), we can always pick a proper b to make sure that the term $\frac{T}{n^{1+\frac{b}{\max\{1,a\}}}}$ in the second bound in Theorem 1 is of order $\frac{1}{n}$. Thus, for all practical purposes, the average regret of DPC when $C = \Theta(n)$ is of order $\sqrt{\frac{\log n}{n}}$. Note that we only need to have a buffer of order $\sqrt{n \cdot \log n}$. Since the magnitude of cumulative demand randomness in n consecutive periods is of order $\sqrt{n}$, this means that we only need to buffer a little bit more (i.e., by a factor of $\sqrt{\log n}$) to guarantee an asymptotically optimal performance under DPC.

REMARK 1 (THE CASE OF SCARCE RESOURCE). Although we have focused our discussions in Theorem 1 on the case $C = \Theta(n)$, the first bound in Theorem 1 also holds when $C = o(n)$. Suppose that demand rates are stationary and $\frac{T}{n} \in \mathcal{Z}^+$. It is not difficult to show in this case that the optimal deterministic solution is static, i.e., $\lambda_t^D = \frac{C}{n}$ for all t. Suppose that $C = n^\gamma$ for some $\gamma \in (0,1)$ and let $\varphi_L = \varphi_U = \frac{1}{2n^{1-\gamma}}$. Then, using $\epsilon = 1 + 6\sqrt{b \cdot n^\gamma \cdot \log n}$ for some $b > 0$ yields an average regret of order $O\left(\sqrt{\frac{b\cdot\log n}{n^{2-\gamma}}}+\frac{T}{n^{1+b}}\right)$. If γ is close to 0 (but not exactly 0), then the average regret of DPC is practically of order $\frac{\sqrt{\log n}}{n}$. This means that DPC has a better performance in the setting of scarce resource. However, there is a caveat: If $C=\Theta(1)$ (e.g., $C=1$), then the argument breaks down and the average regret of DPC is of order $\min\left\{1,\frac{T}{n}\right\}$ (i.e., the performance of DPC can be very poor). This is the setting of an *extremely* scarce resource and a different type of heuristic control seems to be needed to address this case. Since our focus in the paper is on the case $C=\Theta(n)$, we leave this for future research pursuit. (See also Remark 3 at the end of Section 5.)

4.2. Proof of Theorem 1

The proof of Theorem 1 can be separated into two steps. In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t\geq 1$ and $p_t=\hat{p}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC(ϵ).

Step 1

We start with the first step. Let $\Delta_t(\hat{p}_t^D)=D_t(\hat{p}_t^D)-\lambda_t(\hat{p}_t^D)$ (i.e., $\Delta_t(\hat{p}_t^D)$ is the error from the expected demand in period t under price $\hat{p}_t^D$). For notational brevity, we will simply write $\lambda_t=\lambda_t(\hat{p}_t^D)$ and $\Delta_t=\Delta_t(\hat{p}_t^D)$. For some positive $\delta=o(n)$, whose exact value is to be determined later, define a sequence of events $\{\mathcal{A}_k(\epsilon,\delta)\}$ as follows:

$$\mathcal{A}_k(\epsilon,\delta) \;=\; \left\{\max_{t\leq kn}\left|\sum_{s=(k-1)n+1}^{t}\Delta_s\right| \;<\; \delta\right\} \qquad \text{for all } k=1,\ldots,\tfrac{T}{n}. \tag{3}$$

We now analyze $\mathbf{P}(\mathcal{A}_k(\epsilon,\delta))$. Note that, for all $r>0$, we can bound:

$$\begin{aligned}\mathbf{P}\left(\max_{t\leq kn}\left|\sum_{s=(k-1)n+1}^{t}\Delta_s\right| \;\geq\; \delta\right) &\leq \frac{\mathbf{E}\left[\exp\left\{r\left|\sum_{s=(k-1)n+1}^{kn}\Delta_s\right|\right\}\right]}{\exp\{r\delta\}}\\ &\leq \frac{\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]+\mathbf{E}\left[\exp\left\{-r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]}{\exp\{r\delta\}},\end{aligned}$$

where the first inequality follows from a sub-Martingale inequality (see e.g. Williams 1991) and the last inequality holds because $e^{|x|}\leq e^x+e^{-x}$ for all x. Since $D_t(\lambda_t)$ is a Bernoulli random variable with success probability λ_t, by the Moment Generating Function of Bernoulli random variable,

$$\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right] = \prod_{s=(k-1)n+1}^{kn}\mathbf{E}\left[\exp\{r\Delta_s\}\right]$$
$$= \prod_{s=(k-1)n+1}^{kn}\left[e^r\cdot\lambda_t+1-\lambda_t\right]\cdot e^{-r\lambda_t} \leq \prod_{s=(k-1)n+1}^{kn} e^{(e^r-1)\lambda_t}\cdot e^{-r\lambda_t}.$$

Now, for all $|x|\leq 1$, it holds that $e^x-1-x\leq x^2$. Moreover, $\sum_{t=(k-1)n+1}^{kn}\lambda_t = \left(\sum_{t=(k-1)n+1}^{kn}\lambda_t^D\right) - \epsilon \leq \min\{C-\epsilon, n\}$ (because at most one new request arrives in each period). So, we can bound:

$$\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right] \leq \exp\{r^2\min\{C-\epsilon,n\}\} \quad \text{for all } r\in[0,1].$$

Note that similar arguments can also be applied to $\mathbf{E}\left[\exp\left\{-r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]$. Putting all things together, for $r\in[0,1]$, we have:

$$\mathbf{P}(\bar{\mathcal{A}}_k(\epsilon,\delta)) \leq 2\cdot\exp\{r^2\min\{C-\epsilon,n\}-r\delta\} \quad \text{for all } k=1,\dots,\tfrac{T}{n}. \tag{4}$$

Define $\mathcal{G}(\epsilon,\delta) := \cap_{k=1}^{T/n}\mathcal{A}_k(\epsilon,\delta)$. (Per our discussions above, $\mathcal{G}(\epsilon,\delta)$ is our high-probability event.) By the sub-additive property of probability,

$$\mathbf{P}(\mathcal{G}(\epsilon,\delta)) \geq 1-\frac{2T}{n}\exp\{r^2\min\{C-\epsilon,n\}-r\delta\}. \tag{5}$$

We make an important observation— on the set $\mathcal{G}(\epsilon,\delta)$, we always have:

$$\sum_{s=t}^{t+n-1} D_s(\hat{p}_s^D) = \sum_{s=t}^{t+n-1}\left(\lambda_s^D - \frac{\epsilon}{n} + \Delta_s\right) \leq C-\epsilon+3\delta \quad \text{for all } t+n-1\leq T. \tag{6}$$

To see why, note that for any pair (t_1,t_2) with $t_1\in[(k-1)n+1,kn]$ and $t_2\in[kn+1,(k+1)n]$ for some $k\in\{1,\dots,\frac{T}{n}\}$, we have: $\left|\sum_{s=t_1}^{t_2}\Delta_s\right| \leq \left|\sum_{s=t_1}^{kn}\Delta_s\right| + \left|\sum_{s=kn+1}^{t_2}\Delta_s\right| \leq 2\delta+\delta = 3\delta$, where the last inequality follows from the definition of δ in (3). This observation has an important implication: If we set $\delta = \frac{\epsilon-1}{3}$, then we always have $C_t\geq 1$ and $p_t = \hat{p}_t^D$ for all t on the set $\mathcal{G}(\epsilon,\delta)$. For the remaining of the proof, we will therefore assume that $\delta = \frac{\epsilon-1}{3}$.

Step 2

We are now ready to bound the expected regret of DPC(ϵ). Let $\{p_t\}$ be the price sequence under DPC(ϵ) and let $r^u = \max_t \max_{\lambda_t \in \Omega_\lambda} r_t(\lambda_t)$. Note that

$$\begin{aligned}\mathbf{E}[R^{DPC(\epsilon)}] &= \mathbf{E}\left[\sum_{t=1}^{T} r_t(p_t)\right] = \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{p}_t^D)\right)\cdot \mathbf{1}\{\mathcal{G}(\epsilon,\delta)\}\right] \\ &= \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{p}_t^D)\right] - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\tilde{p}_t^D)\right)\cdot \mathbf{1}\{\bar{\mathcal{G}}(\epsilon,\delta)\}\right] \geq \sum_{t=1}^{T} r_t(\tilde{p}_t^D) - r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta)).\end{aligned}$$

Since $J^D - \sum_{t=1}^{T} r_t(\hat{p}_t^D) = \sum_{t=1}^{T} [r_t(p_t^D) - r_t(\hat{p}_t^D)] = \sum_{t=1}^{T} \left[r_t(\lambda_t^D) - r_t(\lambda_t^D - \frac{\epsilon}{n})\right] \leq T \cdot \frac{\Psi\epsilon}{n}$ (because $\epsilon \in [1, n\varphi_L]$, which implies $\lambda_t^D - \frac{\epsilon}{n} \in [\lambda_t^D - \varphi_L, \lambda_t^D + \varphi_U]$, and by Assumption A6), together with the bound in (5) and Assumption A6, we have for all $r \in [0,1]$:

$$\frac{J^D - \mathbf{E}[R^{DPC(\epsilon)}]}{T} \leq \frac{1}{T}\cdot\left[\frac{T\Psi\epsilon}{n} + r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta))\right] \leq \frac{\Psi\epsilon}{n} + \frac{2r^u T}{n}\cdot \exp\{r^2 \min\{C-\epsilon, n\} - r\delta\}$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon, n\}}$ and substituting $\delta = \frac{\epsilon - 1}{3}$ yields:

$$\frac{J^D - \mathbf{E}[R^{DPC(\epsilon)}]}{T} \leq M_1 \cdot \left[\frac{\epsilon}{n} + \frac{T}{n}\cdot \exp\left\{-\frac{(\epsilon-1)^2}{36\min\{C-\epsilon, n\}}\right\}\right] \tag{7}$$

for some $M_1 > 0$ independent of T, C, n, and $\epsilon \in [1, n\varphi_L]$. This completes the proof. ■

5. Deterministic Price Control with Periodic Batch Adjustments

We now discuss an improvement of DPC with periodic batch adjustments. We first provide a description of our heuristic control and then we analyze its performance.

5.1. Control Description and Statement of Result

Let m be a positive integer such that $\frac{n}{m} \in \mathcal{Z}^+$. (This is only exposition clarity and does not affect the key result of our analysis; we discuss this in more detail in Remark 2 at the end of this subsection.) The idea behind our periodic adjustments is to slice the interval $[1, T]$ into $\frac{T}{m}$ batches, each of length m periods, and then to adjust the prices in each batch in such a way that the cumulative errors in the previous batch is corrected in the current batch. To be precise, let $\{\mathcal{T}_i\}_{i=1}^{T/m}$ denote a partition of $[1, T]$, where $\mathcal{T}_i = [(i-1)m+1, im]$ for all $i \geq 1$. For convenience, we assume that

$\mathcal{T}_0 = \emptyset$. Define $\Delta_t(p_t) = D_t(p_t) - \lambda_t(p_t)$ (i.e., $\Delta_t(p_t)$ is the error from expected demand during period t under price p_t), where for notational brevity we will simply write $\Delta_t = \Delta_t(p_t)$. The complete definition of DPC with periodic batch adjustment (DPC-Batch) is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET** and get λ^D.

Step 2. At the beginning of each t, if $t \in \mathcal{T}_i$, do:

a. If $C_t \geq 1$ and $\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}} \Delta_s \in \Omega_\lambda$, set $p_t = \hat{p}_t^D$ where

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}} \Delta_s;$$

b. Otherwise, set $p_t = \bar{p}$.

Unlike the original DPC in Section 4, DPC-Batch is parameterized by two parameters m and ϵ. The value of these parameters must be carefully chosen. If m is too small, the price adjustment scheme under DPC-Batch may not have sufficient corrective power for re-balancing total demands in the current batch (e.g., cumulative errors in the previous batch may have the same order of magnitude as total potential demands in the current batch); if, on the other hand, m is too large, we already incur a lot of loss in the previous batch that is not recoverable by the adjustment in the current batch. The following theorem tells us the performance of DPC-Batch.

THEOREM 2. *Suppose that $\epsilon \in \left[1, \min\left\{n, m, n\cdot\frac{1+4m\cdot\min\{\varphi_L,\varphi_U\}}{4m+n}\right\}\right]$. There exists a constant $M_2 > 0$ such that for all T, C, n, m, and ϵ we have*

$$\text{AvReg}(\textit{DPC-Batch}) \leq M_2 \cdot \left[\frac{\epsilon}{n} + \frac{1}{m} + \frac{T}{m}\cdot\exp\left\{-\frac{(\epsilon-1)^2}{64\min\{C-\epsilon, m\}}\right\}\right]. \tag{8}$$

In particular, if $C = a\cdot n$ for some $a > 0$, then using $\epsilon = 1 + 8\sqrt{b\cdot n^c\cdot\log n}$ and $m = \lceil n^c\rceil$ for some $b > 0$ and $c \in \left(\frac{\log\log n}{\log n}, 1\right)$ yields

$$\text{AvReg}(\textit{DPC-Batch}) = O\left(\frac{\sqrt{b\cdot\log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}\right). \tag{9}$$

Similar to bound (2) in Theorem 1, as long as T grows polynomially in n, we can always pick a proper b such that the term $\frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}$ is of order $\frac{1}{n}$. Thus, the performance of DPC-Batch when

$C = \Theta(n)$ is largely affected by the choice of c. If c is too large (i.e., close to 1), then the bound is again of order $\sqrt{\frac{\log n}{n}}$ as in Theorem 1 (i.e., we do not get any benefit from batch adjustments); if, on the other hand, c is too small (i.e., close to 0), then the bound is of order 1. This means that, under our proposed periodic batch adjustment scheme, the length of each batch m should neither be too small nor too large for the most effective adjustment. Ignoring the logarithmic term in (9), the optimal bound is achieved when $c = 2/3$, which yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. This is a significant improvement over the bound in Theorem 1.

Remark 2 (The Case $\frac{n}{m} \notin \mathcal{Z}^+$). In the proof of Theorem 2, we assume that n is divisible by m for some $m > 1$. If, however, such m does not exist (i.e., n is a prime number), we only need to make a minor change in the definition of a batch. Formally, let $\mathcal{T}_i = [(i-1)m + 1, im]$ for all $i = 1, \ldots, \lfloor \frac{T}{m} \rfloor - 1$, and $\mathcal{T}_{\lfloor \frac{T}{m} \rfloor} = [(\lfloor \frac{T}{m} \rfloor - 1)m + 1, T]$. Note that each of the first $\lfloor \frac{T}{m} \rfloor - 1$ batches still has the same length m, but the length of the last batch is between m and $2m$. With these new batches, the definition of $\hat{p}_t^D$ in Step 2 part a is re-defined as:

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{\epsilon}{n} - \frac{1}{|\mathcal{T}_i|} \sum_{s \in \mathcal{T}_{i-1}} \Delta_s.$$

Following the same arguments as in the proof of Theorem 2 (in Section 5.2), it is not difficult to check that the statement in Theorem 2 still holds under this minor alteration.

Remark 3 (The Case of Scarce Resource). Continuing our discussions in Remark 1 at the end of Section 4, if $C = n^\gamma$ for some $\gamma \in (0,1)$, then using $\epsilon = 1 + 8\sqrt{b \cdot n^{\min\{\gamma,c\}} \cdot \log n}$ yields an average regret of order $O\left(\frac{\sqrt{\log n}}{n^{1-\frac{\min\{\gamma,c\}}{2}}} + \frac{1}{n^c} + \frac{T}{n^{b+\min\{\gamma,c\}}}\right)$. Note that if γ is close to 0 (but not 0), we can choose c close to 1 and the average regret of DPC-Batch is practically of order $\frac{\sqrt{\log n}}{n}$, which is about the same order as the average regret of DPC. This means that, when resource is very scarce, periodic adjustment may not have a significant impact in improving performance.

5.2. Proof of Theorem 2

The proof of Theorem 2 follows similar arguments as the proof of Theorem 1. We still proceed in two steps: In the first step, we construct a high-probability event $\mathcal{G}$ and show that, on the set $\mathcal{G}$, we always have $C_t \geq 1$ and $p_t = \hat{p}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch(m, ϵ).

Step 1

We start with the first step. For some positive $\delta = o(m)$, whose exact value is to be determined later, define a sequence of events $\{\mathcal{A}_i(\epsilon, \delta)\}$ as follows:

$$\mathcal{A}_i(\epsilon,\delta) \;=\; \left\{ \max_{t\leq im} \left| \sum_{s=(i-1)m+1}^{t} \Delta_s \right| \;<\; \delta \right\} \qquad \text{for all } i = 1, \ldots, \tfrac{T}{m}. \tag{10}$$

Analogous to (4) in Section 4.2, it can be shown that

$$\mathbf{P}(\bar{\mathcal{A}}_i(\epsilon,\delta)) \leq 2 \cdot \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\} \quad \text{for all } i = 1, \ldots, \tfrac{T}{m} \text{ and } r \in [0,1]. \tag{11}$$

Now, define $\mathcal{G}(\epsilon,\delta) = \cap_{i=1}^{T/m} \mathcal{A}_i(\epsilon,\delta)$. By the sub-additivity property of probability,

$$\mathbf{P}(\mathcal{G}(\epsilon,\delta)) \geq 1 - \frac{2T}{m} \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\}. \tag{12}$$

We make some important observations. First, on the set $\mathcal{G}(\epsilon,\delta)$, we always have: $\left|\frac{\epsilon}{n} + \frac{1}{m}\sum_{s\in\mathcal{T}_i} \Delta_s\right|$ $\leq \frac{\epsilon}{n} + \frac{\delta}{m}$ for all i. This means that, as long as the parameters ϵ, δ, and m are chosen such that $\frac{\epsilon}{n} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$, the condition $\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}} \Delta_s \in \Omega_\lambda$ in Step 2 part a in the definition of DPC-Batch is always satisfied. For the remaining of the proof, we will therefore assume that $\frac{\epsilon}{n} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$. Now, suppose that $t \in \mathcal{T}_{j_1}$ and $t+n-1 \in \mathcal{T}_{j_2}$, where $j_1 < j_2$ and $t+n-1 \leq T$. We can write the total demands during $[t,\, t+n-1]$ as follows:

$$\begin{aligned}
\sum_{s=t}^{t+n-1} D_s(\hat{p}_s^D) &= \sum_{s\geq t, s\in\mathcal{T}_{j_1}} D_s(\hat{p}_s^D) \;+\; \sum_{j=j_1+1}^{j_2-1}\sum_{s\in\mathcal{T}_j} D_s(\hat{p}_s^D) \;+\; \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} D_s(\hat{p}_s^D) \\
&= \sum_{s\geq t, s\in\mathcal{T}_{j_1}} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j_1-1}} \Delta_l + \Delta_s \right) + \sum_{j=j_1+1}^{j_2-1}\sum_{s\in\mathcal{T}_j} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j-1}} \Delta_l + \Delta_s \right) \\
&\qquad + \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j_2-1}} \Delta_l + \Delta_s \right) \\
&= \sum_{s=t}^{t+n-1} \lambda_s^D - \epsilon - \frac{1}{m}\sum_{s\geq t, s\in\mathcal{T}_{j_1}}\sum_{l\in\mathcal{T}_{j_1-1}} \Delta_l - \sum_{s<t, s\in\mathcal{T}_{j_1}} \Delta_s \\
&\qquad + \left(\sum_{s\in\mathcal{T}_{j_2-1}} \Delta_s - \frac{1}{m}\sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}}\sum_{s\in\mathcal{T}_{j_2-1}} \Delta_s \right) + \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} \Delta_s.
\end{aligned}$$

Since $\mathcal{T}_j$ contains m periods for all j, on the set $\mathcal{G}(\epsilon,\delta)$, we can bound:

$$\left|\frac{1}{m}\sum_{s\geq t, s\in\mathcal{T}_{j_1}}\sum_{l\in\mathcal{T}_{j_1-1}}\Delta_l\right|\leq\delta,\quad \left|\sum_{s<t, s\in\mathcal{T}_{j_1-1}}\Delta_s\right|\leq\delta,\quad \left|\sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}}\Delta_s\right|\leq\delta$$

$$\text{and}\quad \left|\sum_{s\in\mathcal{T}_{j_2-1}}\Delta_s-\frac{1}{m}\sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}}\sum_{s\in\mathcal{T}_{j_2-1}}\Delta_s\right|\leq\delta.$$

Putting the above four bounds together, on the set $\mathcal{G}(\epsilon,\delta)$, we have:

$$\sum_{s=t}^{t+n-1}D_s(\hat{p}_s^D)\leq C-\epsilon+4\delta\quad\text{for all } t+n-1\leq T. \tag{13}$$

It is worth noting that although (13) is similar to (6) in the proof of Theorem 1, the term δ in (6) represents a bound on cumulative errors during n periods whereas the term δ in (13) represents a bound on cumulative errors during $m<n$ periods (i.e., the δ in (13) is potentially much smaller than the δ in (6), which highlights the potential improvement due to batch adjustments).

Let $\delta=\frac{\epsilon-1}{4}$. Given this and the assumption that $\frac{\epsilon}{n}+\frac{\delta}{m}\leq\min\{\varphi_L,\varphi_U\}$, it is not difficult to see that the following always hold on $\mathcal{G}(\epsilon,\delta)$: $C_t\geq 1$ and $\lambda_t^D-\frac{\epsilon}{n}-\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\in\Omega_\lambda$ for all i and $t\in\mathcal{T}_i$. As a consequence, we also have $p_t=\hat{p}_t^D$ for all t.

Step 2

We are now ready to bound the expected regret of DPC-Batch(m,ϵ). Let $\{p_t\}$ be the price sequence under DPC-Batch(m,ϵ). Note that

$$\begin{aligned}\mathbf{E}[R^{DPC-Batch(m,\epsilon)}]&=\mathbf{E}\left[\sum_{t=1}^T r_t(p_t)\right]\geq\mathbf{E}\left[\left(\sum_{t=1}^T r_t(\hat{p}_t^D)\right)\cdot\mathbf{1}\{\mathcal{G}(\epsilon,\delta)\}\right]\\&=\mathbf{E}\left[\sum_{t=1}^T r_t(\hat{p}_t^D)\right]-\mathbf{E}\left[\left(\sum_{t=1}^T r_t(\hat{p}_t^D)\right)\cdot\mathbf{1}\{\bar{\mathcal{G}}(\epsilon,\delta)\}\right].\end{aligned}$$

The second expectation after the last equality above can be bounded by $r^uT\cdot\mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta))$ where $r^u=\max_t\max_{\lambda_t\in\Omega_\lambda}r_t(\lambda_t)$. As for the first expectation, suppose that $t\in\mathcal{T}_i$ for some $i\geq 2$. By Taylor's expansion and Assumption A6, we can bound $r_t(\hat{p}_t^D)=r_t\left(\lambda_t^D-\frac{\epsilon}{n}-\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right)\geq r_t(\lambda_t^D)-r_t'(\lambda_t^D)\cdot\left(\frac{\epsilon}{n}+\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right)-\Psi\cdot\left(\frac{\epsilon}{n}+\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right)^2$. Taking expectation and applying Assumption A6 one more time yield $\mathbf{E}\left[r_t(\hat{p}_t^D)\right]\geq r_t(\lambda_t^D)-\frac{\Psi\epsilon}{n}-\Psi\cdot\left(\frac{2\epsilon^2}{n^2}+\frac{2}{m}\right)$, where the inequality follows because $(x+y)^2 \le 2x^2 + 2y^2$ for all (x,y) and $\mathbf{E}\left[\left(\sum_{s\in\mathcal{T}_{i-1}} \Delta_s\right)^2\right] \le m$ (by definition, $\{\Delta_s\}_{s\in\mathcal{T}_{i-1}}$ are independent zero-mean random variables and $|\Delta_s| \le 1$).

Putting the bounds together, for all $r \in [0,1]$, we have:

$$\begin{aligned}\frac{J^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T} &\le \frac{1}{T}\cdot\left[\frac{T\Psi\epsilon}{n} + T\Psi\cdot\left(\frac{2\epsilon^2}{n^2}+\frac{2}{m}\right) + r^u T\cdot P(\bar{\mathcal{G}}(\epsilon,\delta))\right] \\ &\le \frac{\Psi\epsilon}{n} + \frac{2\Psi\epsilon^2}{n^2} + \frac{2\Psi}{m} + \frac{2r^u T}{m}\cdot\exp\{r^2\min\{C-\epsilon, m\} - r\delta\}.\end{aligned}$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon,m\}}$ and substituting $\delta = \frac{\epsilon-1}{4}$ yields:

$$\frac{J^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T} \le M_2\cdot\left[\frac{\epsilon}{n} + \frac{1}{m} + \frac{T}{m}\cdot\exp\left\{-\frac{(\epsilon-1)^2}{64\min\{C-\epsilon,m\}}\right\}\right] \tag{14}$$

for some $M_2 > 0$ independent of T, C, n, m, and $\epsilon \in \left[1, \min\left\{n, m, n\cdot\frac{1+4m\cdot\min\{\varphi_L,\varphi_U\}}{4m+n}\right\}\right]$. (Note that $\delta = \frac{\epsilon-1}{4}$ and $\epsilon \le \frac{n+4mn\cdot\min\{\varphi_L,\varphi_U\}}{4m+n}$ implies $\frac{\epsilon}{n} + \frac{\delta}{m} \le \min\{\varphi_L,\varphi_U\}$; $1 < \epsilon < m$ ensures that $r = \frac{\delta}{2\min\{C-\epsilon,m\}} = \frac{\epsilon-1}{8\min\{C-\epsilon,m\}} \in (0,1)$.) To get bound (9), we further require $c > \frac{\log\log n}{\log n}$ to ensure $r \in (0,1)$. This completes the proof. ■

6. Numerical Experiments

We now conduct simple numerical experiments to illustrate the performance of the proposed heuristic controls under different problem parameters. For simplicity, we assume that the demand function (i.e., purchasing probability) is stationary over time, and is exponentially decreasing in price, i.e., $\lambda(p) = \exp\left(\lambda_0 - \frac{p}{p_0}\right)$. We use $\lambda_0 = 0.8$ and $p_0 = 100$. The length of selling horizon T and resource capacity C are both set to be linear in n, and we vary n from $1,000$ to $8,000$. Specifically, we choose $C = 0.7\cdot n$ and $T = 5\cdot n$ (i.e., $n = 1,000$ corresponds to the problem instance with 700 units of resources and length of selling horizon equals 5,000 periods). The resulting deterministic problem has a stationary optimal solution $\lambda_t^D \equiv \lambda^D = 0.7$, with optimal objective value $J^D = 80.97$.

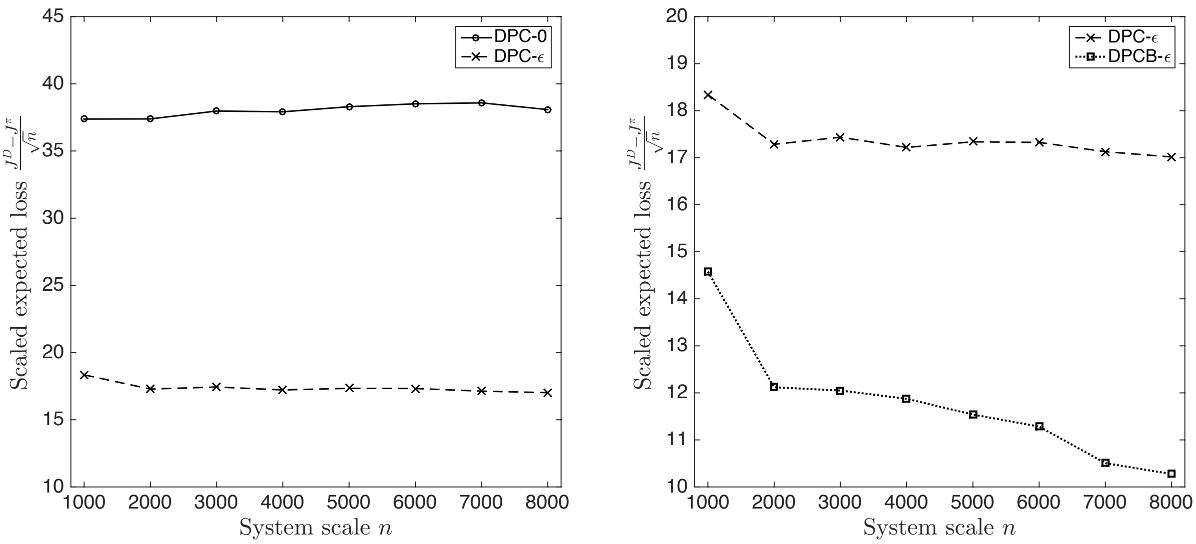
Table 1 summarizes the heuristic controls tested in all experiments. Several comments are in order regarding the implementation details. First, we implement DPC-0 to identify the impact of injecting a buffer for deterministic control. Second, in DPCB-ϵ, if $T/m \notin \mathbb{Z}_+$, we simply set the last period to be of size T mod m (given the discussions in Remark 2 in Section 5, this should not affect the performance of DPC-Batch by much). Third, LRC-k refers to a simple adoption of the self-adjusting control proposed in Jasin (2014), where we simply re-start the control at every k periods by setting the cumulative error to be zero (see Section EC.4 for a detailed description). Fourth, for any combination of parameters, we simulate all the heuristic controls with 300 Monte Carlo runs to approximate their expected total revenues. Lastly, for DPC-ϵ and DPCB-ϵ, we simply use a grid-search method to find the optimal ϵ (between 0 and 1, with an increment of 0.01).

Table 1 Summary description of all heuristic controls

Label	Description
DPC-0	DPC(ϵ) with $\epsilon = 0$ (defined in Section 4.1)
DPC-ϵ	DPC(ϵ) (defined in Section 4.1)
DPCB-ϵ	DPC-Batch(m, ϵ) with $m = \lceil n^{2/3} \rceil$ (defined in Section 4.1)
LRC-k	Linear rate control with re-starting at every k periods (see Section 4 in Jasin 2014)

Simulation results. Figures 1 and 1 show the expected total regrets of the first three heuristic controls, where the y-axis is the *scaled* total expected revenue loss $\frac{J^D - J^\pi}{\sqrt{n}}$. We do not plot the regret of LRC-k since it performs much worse than any of the other three heuristic controls under any k, but we report the complete numerical results in Section EC.4 in the electronic companion. As expected, DPC-Batch dominates DPC, which in turn dominates DPC-0. Moreover, a closer look on the average regret confirms the asymptotic optimality of all three heuristic controls. The relatively poor performance (in fact, may not even be asymptotically optimal) of LRC suggests that a heuristic control that performs well in the setting of canonical revenue management cannot be directly adopted to the setting of revenue management with reusable resources and deterministic service time requirements. This reinforces our point in Section 1 that the setting considered in our work, though may appear identical, is not exactly the same as the setting in the standard revenue management literature.

Figure 1 Plot of expected average regret (%) with different ϵ



7. Extension to Multiple Service Types with Heterogeneous Service Time Requirements

In this section, we discuss a generalization of the basic model in Section 3 that allows different service types with heterogeneous service time requirements. We first discuss the setting of the problem and then provide a generalization of DPC-Batch.

7.1. The Setting

The firm sells $K \geq 1$ service (or product) types where a request of service type k requires one unit of resource and n_k units of service time (or n_k periods). For ease of exposition, we will assume that $\frac{T}{n_k} \in \mathcal{Z}^+$ for all $k = 1, \ldots, K$. Moreover, without loss of generality, we also assume that the service types are labeled in such a way that $1 \leq n_1 \leq n_2 \leq \cdots \leq n_K$. The dynamics of the problems are as follows: At the beginning of period t, the firm sets the prices for all service types, denoted by a vector $\mathbf{p}_t := (p_{t,1}, \ldots, p_{t,K}) \in \Omega_p$. (Unless otherwise noted, all vectors are to be understood as column vectors.) For period t, a price vector $\mathbf{p}_t$ induces a demand vector $\mathbf{D}_t(\mathbf{p}_t) = (D_{t,1}(\mathbf{p}_t), \ldots, D_{t,K}(\mathbf{p}_t))$ with rate vector $\boldsymbol{\lambda}_t(\mathbf{p}_t) := (\lambda_{t,1}(\mathbf{p}_t), \ldots, \lambda_{t,K}(\mathbf{p}_t))$, where $\boldsymbol{\lambda}_t(\mathbf{p}_t) = \mathbf{E}[\mathbf{D}_t(\mathbf{p}_t)]$. The corresponding revenue rate is given by $r_t(\mathbf{p}_t) = \mathbf{E}[\mathbf{p}_t^\top \mathbf{D}_t(\mathbf{p}_t)] = \mathbf{p}_t^\top \boldsymbol{\lambda}_t(\mathbf{p}_t)$. By the invertibility assumption (see below), we will also use $\mathbf{D}_t(\boldsymbol{\lambda}_t) = \mathbf{D}_t(\mathbf{p}_t(\boldsymbol{\lambda}_t))$ and $r_t(\boldsymbol{\lambda_t}) = \boldsymbol{\lambda}_t^\top \mathbf{p}_t(\boldsymbol{\lambda}_t)$ to denote the realized demand vector and revenue rate as a function of demand rates, respectively. As in the basic model, we assume that demands across different periods are independent but demands over different service types within the same period may be correlated and demand rates as functions of time may be non-stationary. We assume at most one request arrives in each period, i.e., $\sum_{k=1}^K D_{t,k}(\mathbf{p}_t) \le 1$ (this is without loss of generality). Let $\Omega_p = \otimes_{k=1}^K \Omega_{p,k}$ and $\Omega_\lambda = \otimes_{k=1}^K \Omega_{\lambda,k}$ denote the convex feasible set for price vector and demand rate vector, respectively. The following regularity conditions are the generalization of Assumptions A1-A4 in Section 3 to the multiple service types setting:

MA1. $\boldsymbol{\lambda}_t(\mathbf{p}_t) : \Omega_p \to \Omega_\lambda$ is bounded, twice differentiable and invertible.

MA2. For each k, there exists a "turn-off" price $\bar{p}_k$ such that $p_{k,t}^v \to \bar{p}_k$ implies $\lambda_{t,k}(\mathbf{p}_t^v) \to 0$.

MA3. $\boldsymbol{\lambda}_t^v \to \mathbf{0}$ implies $r_t(\boldsymbol{\lambda}_t^k) \to 0$ for any feasible sequence $\{\boldsymbol{\lambda}_t^v\}_{v=1}^\infty$.

MA4. $r_t(\boldsymbol{\lambda}_t)$ is bounded, strictly jointly concave, and has a finite maximizer $\boldsymbol{\lambda}_t^* \in \Omega_\lambda$.

The optimal stochastic control formulation of our dynamic pricing problem is given by:

$$\textbf{OPT-M}: \quad J_M^* = \left\{ \max_{\pi \in \Pi} \mathbf{E}\left[\sum_{t=1}^T r_t(\mathbf{p}_t^\pi)\right] : \sum_{k=1}^K \sum_{s=\max\{1,\, t-n_k+1\}}^t D_{s,k}(\mathbf{p}_s^\pi) \le C \ \text{ for all } t \le T \right\}$$

where the constraints must hold almost surely (or with probability one) and Π is the set of all non-anticipating controls. Using demand rate vector as the decision variable, the deterministic relaxation of **OPT-M** is given by:

$$\textbf{DET-M}: \quad J_M^D = \left\{ \max_{\boldsymbol{\lambda}_t \in \Omega_\lambda} \sum_{t=1}^T r_t(\boldsymbol{\lambda}_t) : \sum_{k=1}^K \sum_{s=\max\{1,\, t-n_k+1\}}^t \lambda_{s,k} \le C \ \text{ for all } t \le T \right\}$$

As in Lemma 1, it is not difficult to show that J_M^D is an upper bound of J_M^*. Therefore, the average regret defined in Section 3 can still be used as a proper performance measure. Let $\boldsymbol{\lambda}^D := (\boldsymbol{\lambda}_t^D)_{t=1}^T$ denote the optimal solution of **DET-M**, and let $\mathbf{p}^D := (\mathbf{p}_t^D)_{t=1}^T$ denote the corresponding optimal price vectors (i.e., $\mathbf{p}_t^D = \mathbf{p}_t(\boldsymbol{\lambda}_t^D)$). Let $\mathbf{e}$ be a vector of ones, with a proper dimension. Similar to Assumptions A5-A6, we assume that there exist positive constants φ_L , φ_U, and Ψ such that the following two conditions hold for all t:

MA5. $[\boldsymbol{\lambda}_t^D - \varphi_L \mathbf{e}, \boldsymbol{\lambda}_t^D + \varphi_U \mathbf{e}] \subseteq \Omega_\lambda$.

MA6. $||\nabla r_t(\boldsymbol{\lambda})||_\infty$ and $||\nabla^2 r_t(\boldsymbol{\lambda})||_2$ are bounded from above by Ψ on $[\lambda_t^D - \varphi_L, \lambda_t^D + \varphi_U]$.

We are now ready to present the generalization of DPC-Batch in the setting with multiple service types and heterogeneous service time requirements.

7.2. A Generalized DPC-Batch and Its Performance

Let $\mathbf{m} = (m_1, \ldots, m_K)$ be a sequence of positive integers such that $\frac{n_k}{m_k} \in \mathcal{Z}^+$ for all k. (As in Section 5.1, the existence of such sequence is assumed for ease of exposition and does not affect our result. If a proper m_k satisfying $\frac{n_k}{m_k} \in \mathcal{Z}^+$ does not exist, then we can slightly modify our batch definition as in Remark 2 at the end of Section 5.1.) For each service type k, we slice the selling horizon into $\frac{T}{m_k}$ batches, each of length m_k periods. Let $\mathcal{T}_{k,i} = [(i-1)m_k + 1, im_k]$ denote the i^{th} batch for service type k. The key idea behind our generalized DPC-Batch is to manage the demand rate for each service type somewhat independently of the other service types. To be precise, the demand rates in each batch are adjusted in such a way that the cumulative errors for a given service type in the previous batch are corrected by the demands of the same service type in the current batch. (This does not mean that the controls are completely decoupled since demands over different service types are still connected through their prices, which means that the corresponding prices adjustments need to be computed jointly.) Let $\boldsymbol{\Delta}_t := (\Delta_{t,k})_{k=1}^K = (D_{t,k}(\mathbf{p}_t) - \lambda_{t,k}(\mathbf{p}_t))_{k=1}^K$ denote the vector of errors from expected demands in period t under price vector $\mathbf{p}_t$ (we suppress the dependency of $\boldsymbol{\Delta}_t$ on $\mathbf{p}_t$). Also, let $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_K)$ be a sequence of real-valued constants denoting the size of buffer for each service type, and define $i_k(t)$ such that $t \in \mathcal{T}_{k,i_k(t)}$ for all t and k. The complete definition of our generalized DPC-Batch with multiple service types and heterogeneous service time requirements is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET-M** and get $\boldsymbol{\lambda}^D$.

Step 2. At the beginning of each t, do:

a. If $C_t \geq 1$ and $\lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k} \in \Omega_{k,\lambda}$, set $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ where

$$\lambda_{t,k}(\hat{\mathbf{p}}_t^D) \;=\; \lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k} \quad \text{for all } k;$$

b. Otherwise, set $\mathbf{p}_t = \bar{\mathbf{p}}$.

Note that the price vector $\hat{\mathbf{p}}_t^D$ in Step 2 part a is well-defined by the invertibility assumption in MA1. Let $C_k^D := \max_{1\leq t\leq T}\sum_{s=\max\{1,\,t-n_k+1\}}^{t} \lambda_{s,k}^D$ denote the maximum amount of resource used by service type k in the deterministic model. The following theorem tells us the performance of DPC-Batch with heterogeneous service time requests; we defer its proof to the electronic companion.

Theorem 3. *Suppose that $0<n_1\leq\cdots\leq n_K\leq 1$. There exists a constant $M_3>0$ such that for all T, C, m_k, $\epsilon_k \in n_k\cdot\left[\frac{1}{Kn_1}, \min\left\{1, \frac{1}{K}\cdot\frac{1+4Km_k\cdot\min\{\varphi_L,\varphi_U\}}{4m_k+n_1}\right\}\right]$, and $n_1 \geq \frac{1}{K\min\{\varphi_L,\varphi_U\}}$ we have*

$$\textsc{AvReg}(\textit{DPC-Batch}) \leq M_3\cdot\sum_{k=1}^{K}\left[\frac{\epsilon_k}{n_k} + \frac{1}{m_k} + \frac{T}{m_k}\cdot\exp\left\{-\frac{(Kn_1\epsilon_k - n_k)^2}{64K^2n_k^2\min\{C_k^D-\epsilon_k,\, m_k\}}\right\}\right]. \quad (15)$$

In particular, if $n_k=\alpha_k\cdot n$ and $C_k^D=\beta_k\cdot n$ for some $0<\alpha_1\leq\cdots\leq\alpha_K$ and $\beta_k>0$ for all k, then using $\epsilon_k=\frac{\alpha_k}{K\alpha_1}(1+8\sqrt{b\cdot n^c\cdot\log n})$ and $m_k=\lceil n^c\rceil$ for all k, for some $b>0$ and $c\in[0,1)$, yields

$$\textsc{AvReg}(\textit{DPC-Batch}) = O\left(\frac{\sqrt{b\cdot n^c\cdot\log n}}{n} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,\max_k\{\beta_k\}\}}}}\right). \quad (16)$$

Two comments are in order. First, under a proper choice of b, setting $c=2/3$ in (16) yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. This is the same order as the optimal bound as in Theorem 2 (with $c=2/3$). Second, although the second bound in Theorem 3 only focuses on the case where $C_k^D=\Theta(n_k)=\Theta(n)$ for all k, the first bound in Theorem 3 holds in great generality. For example, if $n_k=\Theta(n^{\alpha_k})$ and $C_k^D=\beta_k\cdot n_k$ for some $\alpha_k,\beta_k>0$, we can use $\epsilon_k=\frac{n_k}{Kn_1}\left(1+8K\sqrt{b\cdot n_k^{c_k}\cdot\log n_k}\right)$ and $m_k=\lceil n_k^{c_k}\rceil$ for some $c_k>0$ for all k, for some $b>0$, and the bound in Theorem 3 becomes

$$\textsc{AvReg}(\textit{DPC-Batch}) = \sum_{k=1}^{K} O\left(\frac{\sqrt{b\cdot n_k^{c_k}\cdot\log n_k}}{n_1} + \frac{1}{n_k^{c_k}} + \frac{T}{n_k^{c_k+\frac{b}{\max\{1,\beta_k\}}}}\right).$$

Ignoring the logarithmic term in the bound above, an optimal c_k can be calculated by setting $n_k^{\frac{3}{2}c_k} = n_1$, or equivalently $c_k = \frac{2}{3} \cdot \frac{\log n_1}{\log n_k} := c_k^*$. Note that the number of batches in one service cycle for service type k under $c_k = c_k^*$ is approximately $n_k^{1-c_k^*} = n_1^{\frac{2}{3}(\frac{1}{c_k^*}-1)}$. Since the power term on n_1 is decreasing in c_k^* for all $c_k^* \in (0,1)$ and a larger n_k implies a smaller c_k^*, the service type with a longer service time requires a larger batch size and a more frequent price adjustments during one service cycle than the service type with a shorter service time. Overall, the average regret of DPC-Batch for the above scenario under c_k^* is of order $\frac{\sqrt{\log n_1}}{n_1^{2/3}}$.

8. Extension to Advance Service Bookings with Homogeneous Service Time Requirements

In this section, we consider a generalization of the basic model in Section 3 to the setting with advance service booking or scheduling. We first discuss the setting of the problem and then we provide a generalization of DPC-Batch.

8.1. The Setting

Similar to the basic model, the firm sells only a single service type where each request requires a single unit of resource and n units of service time. However, unlike in the basic model where a customer arriving in period t immediately starts her service in period t, she can now choose to start her service at time $t + \ell$, where $\ell \in [0, L]$. (For simplicity, we will call a request whose service starts ℓ periods later as type-ℓ request; this should not be confused with the meaning of "type" in the previous section.) The firm controls the arrival rates of all types of requests by setting a price vector $\mathbf{p}_t = (p_{t,0}, \ldots, p_{t,L})$, where $p_{t,\ell}$ is the price of a type-ℓ request booked in period t (note that $p_{t,0}$ is the price of service that starts immediately in period t). Demand rates in period t is denoted by $\boldsymbol{\lambda}_t(\mathbf{p}_t) := (\lambda_{t,0}(\mathbf{p}_t), \ldots, \lambda_{t,L}(\mathbf{p}_t))$. Let $\mathbf{D}_t(\mathbf{p}_t) = (D_{t,1}(\mathbf{p}_t), \ldots, D_{t,L}(\mathbf{p}_t))$ denote the realized requests in period t (by definition, $\mathbf{E}[\mathbf{D}_t(\mathbf{p}_t)] = \boldsymbol{\lambda}_t(\mathbf{p}_t)$). By the invertibility assumption (see below), we can write the corresponding revenue rate as $r_t(\mathbf{p}_t) := \mathbf{p}_t^\top \boldsymbol{\lambda}_t(\mathbf{p}_t) = \boldsymbol{\lambda}_t^\top \mathbf{p}_t(\boldsymbol{\lambda}_t) = r_t(\boldsymbol{\lambda}_t)$. As in the basic model, we assume that demands across different periods are independent, though demands over different request types within the same period may be correlated, and at most one request arrives in each period, i.e., $\sum_{\ell=0}^{L} D_{t,\ell}(\mathbf{p}_t) \leq 1$. Let $\Omega_p = \otimes_{\ell=0}^{L}\Omega_{p,\ell}$ and $\Omega_\lambda = \otimes_{\ell=0}^{L}\Omega_{\lambda,\ell}$ denote the convex feasible set for price vector and demand rate vector, respectively. As in Section 3, we assume that MA1-MA4 hold. (Although the definition of service, or request, types in Sections 7 and 8 are different, from the point of view of abstraction, the demand and revenue functions in Sections 7 and 8 are essentially a multi-product variant of the functions in Section 3.)

The optimal stochastic control formulation of our dynamic pricing problem is given by:

$$\textbf{OPT-A}: \quad J_A^* = \left\{ \max_{\pi \in \Pi} \mathbf{E}\left[\sum_{t=1}^{T} r_t(\mathbf{p}_t^\pi)\right] : \sum_{\ell=0}^{L} \sum_{s=\max\{1,t-n-\ell+1\}}^{t-\ell} D_{s,\ell}(\mathbf{p}_s^\pi) \leq C \ \text{ for all } t \leq T \right\}$$

where the constraints must hold almost surely (or with probability one) and Π is the set of all non-anticipating controls. Using demand rate vector as the decision variable, the deterministic relaxation of **OPT-A** is given by:

$$\textbf{DET-A}: \quad J_A^D = \left\{ \max_{\boldsymbol{\lambda}_t \in \Omega_\lambda} \sum_{t=1}^{T} r_t(\boldsymbol{\lambda}_t) : \sum_{\ell=0}^{L} \sum_{s=\max\{1,t-n-\ell+1\}}^{t-\ell} \lambda_{s,\ell} \leq C \ \text{ for all } t \leq T \right\}$$

Let $\boldsymbol{\lambda}^D := (\boldsymbol{\lambda}_t^D)_{t=1}^T$ denote the optimal solution of **DET-A**, and let $\mathbf{p}^D := (\mathbf{p}_t(\boldsymbol{\lambda}_t^D))_{t=1}^T$ denote the corresponding price vectors. As in Section 7, we assume that MA 5 and MA 6 also hold for all t.

Lastly, we define our performance measure in the setting with advance booking as follows:

$$\textsc{AvReg}(\pi) = \frac{J_A^D - \mathbf{E}[R^\pi]}{T \cdot (L+1)}.$$

In the same spirit with Lemma 1, it is not difficult to show that $J_A^* \leq J_A^D$. However, unlike in the basic model in Section 3 where the expected total revenues under the optimal policy throughout T periods only scales linearly with T, the expected total revenues in the advance booking setting may scale linearly with $T \cdot (L+1)$, especially when T is large and the demand rate function $\lambda_{t,\ell}(\cdot)$ has the same order of magnitude for all t and ℓ (i.e., at any period t, we have the same intensity among customers who are requesting to start their service at period $t+\ell$ where $\ell = 0, 1, \ldots, L$), because we are essentially collecting revenues from about $n \cdot (L+1)$ customers instead of n during each service cycle. This explains why we divide the expected total regrets with $T \cdot (L+1)$ instead of T in the above. We can alternatively interpret this as the average expected revenue loss *per* customer.

8.2. A Generalized DPC-Batch and Its Performance

Let $\{\mathcal{T}_i\}_{i=1}^{T/m}$ denote a partition of $[1, T]$, where $\mathcal{T}_i = [(i-1)m+1, im]$ for all $i \geq 1$. The key idea behind our generalized DPC-Batch with advance service booking is to correct the cumulative errors of type-ℓ request in the previous batch with the demands of type-ℓ request in the current batch. Let $\mathbf{\Delta}_t := (\Delta_{t,\ell})_{\ell=0}^{L} = (D_{t,\ell}(\mathbf{p}_t) - \lambda_{t,\ell}(\mathbf{p}_t))_{\ell=0}^{L}$ denote the vector of errors from expected demands in period t under price vector $\mathbf{p}_t$, where we suppress the dependency of $\mathbf{\Delta}_t$ on $\mathbf{p}_t$. For each t, let $i_\ell(t)$ be such that $\max\{t-\ell, 1\} \in \mathcal{T}_{i_\ell(t)}$. The complete definition of DPC-Batch with advance service booking is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET-A** and get $\boldsymbol{\lambda}^D$.

Step 2. At the beginning of each $t \geq 1$, do:

a. If $C_t \geq 1$ and $\lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \in \Omega_{\lambda,\ell}$, set $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ where

$$\lambda_{t,\ell}(\hat{\mathbf{p}}_t^D) \; = \; \lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \quad \text{for all } \ell;$$

b. Otherwise, set $\mathbf{p}_t = \bar{\mathbf{p}}$.

The following theorem tells us the performance of DPC-Batch with advance service booking; we defer its proof to the electronic companion.

THEOREM 4. *The following two bounds hold for all C, L, n and m:*

1. *If $L \leq n$ and $\epsilon \in \left[1, \min\left\{n(L+1), m(L+1), n\cdot\frac{1+m(L+1)\min\{\varphi_L,\varphi_U\}}{8m+n}\right\}\right]$, there exists a constant $M_4 > 0$ such that*

$$\text{AVREG}(\textit{DPC-Batch}) \leq M_4 \cdot \left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m}\cdot\exp\left\{-\frac{(\epsilon-1)^2}{256(L+1)^2\min\{C-\epsilon,\, m\}}\right\}\right] \tag{17}$$

2. *If $L > n$ and $\epsilon \in (L+1)\cdot\left[2, \min\left\{n, m, \frac{4mn\min\{\varphi_L,\varphi_U\}+2}{4m+n}\right\}\right]$, there exists a constant $M_4' > 0$ such that*

$$\text{AVREG}(\textit{DPC-Batch}) \leq M_4' \cdot \left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m}\cdot\exp\left\{-\frac{(\epsilon-2(L+1)/n)^2}{64(L+1)^2\min\{C-\epsilon,\, m\}}\right\}\right]. \tag{18}$$

In particular, if $C = a \cdot nL$, $L = n^d$, $m = \lceil n^c \rceil$ *for some* $a > 0$, $d \geq 0$ *and* $c \in \left(\frac{\log\log n}{\log n}, 1\right)$, *we can bound the average regret of DPC-Batch as follows:*

1. *For* $d \leq 1$, *using* $\epsilon = 1 + 16\sqrt{b \cdot n^{2d+c} \cdot \log n}$ *for some* $b > 0$ *in bound (17) yields*

$$\text{AvReg}(\textit{DPC-Batch}) \;=\; O\left(\frac{\sqrt{b \log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}\right). \tag{19}$$

2. *For* $d > 1$, *using* $\epsilon = 2 \cdot \frac{L+1}{n} + 8\sqrt{b \cdot n^{2d+c} \cdot \log n}$ *for some* $b > 0$ *in bound (18) yields*

$$\text{AvReg}(\textit{DPC-Batch}) \;=\; O\left(\frac{\sqrt{b \cdot \log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}\right). \tag{20}$$

The two general bounds in Theorem 4 (i.e., (17) and (18)) are proved in a very similar manner under different requirements on ϵ and the relative magnitude of n and L. Together, they are the analogue of (8) in Theore 2 and holds for general problem parameters C, m, and n. Under slightly different choice of ϵ, the optimal order of (19) and (20) are both achieved when $c = 2/3$, which yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. Hence, Theorem 4 tells us DPC-Batch can be generalized to the setting with advance service bookings without worsening the performance.

9. Closing Remarks

In this paper, we address the dynamic pricing problem with reusable resources and deterministic service time requirements. Given the complexity of solving the stochastic control optimally, we focus on designing provably-good heuristic controls and evaluate their performances in the asymptotic regime. We also extend our result to the setting with heterogeneous service time requirements and advance booking length. Given its simplicity and generality, we believe that our heuristic controls can be tailored to address practical dynamic pricing problems faced by firms from various industries. Methodologically, our asymptotic analysis also shed lights on the difference between revenue management with reusable resources and deterministic service time requirements and the canonical revenue management problems. Many possible extensions are not addressed in this paper. For example, it is interesting to see how our analytical framework can be generalized to the setting with stochastic service time. Another potential future direction is to analyze the "network" version of our model, where resources can move dynamically between nodes, which is the setting of many on-demand ride sharing models such as UBER and Lyft.

References

Afèche, P, B Ata. 2013. Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *M&SOM* **15**(2) 292–304.

Al-Roomi, M, S Al-Ebrahim, S Buqrais, I Ahmad. 2013. Cloud computing pricing models: a survey. *Int. J. Grid Dist. Comput.* **6**(5) 93–106.

Altman, R, T Jiménez, G Koole. 2001. On optimal call admission control in resource-sharing system. *IEEE T. Commun.* **49**(9) 1659–1668.

Alzhouri, F, A Agarwal. 2015. Dynamic pricing scheme: Towards cloud revenue maximization. *2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom)*. IEEE, 168–173.

Arshad, S, S Ullah, SA Khan, MD Awan, MSH Khayal. 2015. A survey of cloud computing variable pricing models. *2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)*. IEEE, 27–32.

Asdemir, K, VS Jacob, E Krishnan. 2009. Dynamic pricing of multiple home delivery options. *Eur. J. Oper. Res.* **196**(1) 246–257.

Ata, B, TL Olsen. 2009. Near-optimal dynamic lead-time quotation and scheduling under convex-concave customer delay costs. *Oper. Res.* **57**(3) 753–768.

Ata, B, TL Olsen. 2013. Congestion-based leadtime quotation and pricing for revenue maximization with heterogeneous customers. *Queueing Syst.* **73**(1) 35–78.

Ata, B, S Shneorson. 2006. Dynamic control of an m/m/1 service system with adjustable arrival and service rates. *Manage. Sci.* **52**(11) 1778–1791.

Atar, R, MI Reiman. 2012. Asymptotically optimal dynamic pricing for network revenue management. *Stochastic Syst.* **2**(2) 232–276.

Ben-Tal, A, A Goryashko, E Guslitzer, A Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Math. Program.* **99**(2) 351–376.

Bensinger, G. 2015. Startups scramble to define employee. `http://goo.gl/SLLcqP`.

Bertsimas, D, DA Iancu, PA Parrilo. 2010. Optimality of affine policies in multistage robust optimization. *Math. of Oper. Res.* **35**(2) 363–394.

Besbes, O, C Maglaras. 2009. Revenue optimization for a make-to-order queue in an uncertain market environment. *Oper. Res.* **57**(6) 1438–1450.

Bitran, G, R Caldentey. 2003. An overview of pricing models for revenue management. *M&SOM* **5**(3) 203–229.

Borgs, C, O Candogan, J Chayes, I Lobel, H Nazerzadeh. 2014. Optimal multiperiod pricing with service guarantees. *Manage. Sci.* **60**(7) 1792–1811.

Brown, L, N Gans, A Mandelbaum, A Sakov, H Shen, S Zeltyn, L Zhao. 2005. Statistical analysis of a telephone call center: A queueing-science perspective. *J. Ame. Stat. Assoc.* **100**(469) 36–50.

Businessweek, Bloomberg. 2015. Nick kokonas is selling tickets to dinner. `http://goo.gl/0St8EU`.

Calafiore, GC. 2009. An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM J. Control Optim.* **48**(4) 2254–2274.

Campbell, AM, M. Savelsbergh. 2006. Incentive schemes for attended home delivery services. *Transport. Sci.* **40**(3) 327–341.

Çelik, S., C. Maglaras. 2008. Dynamic pricing and lead-time quotation for a multiclass make-to-order queue. *Manage. Sci.* **54**(6) 1132–1146.

Chamlee, V. 2016. Why surge pricing could hit food delivery apps next. `http://goo.gl/qbv80S`.

Chen, H, MZ Frank. 2001. State dependent pricing with a queue. *IIE Trans.* **33**(10) 847–860.

Chen, Q, S Jasin, I Duenyas. 2014. Adaptive parametric and nonparametric multi-product pricing via self-adjusting controls. Draft available at SSRN 2533468.

Chen, Q, S Jasin, I Duenyas. 2015. Real-time dynamic pricing with minimal and flexible price adjustment. *Manage. Sci* .

Chen, Y, VF Farias. 2016. Robust dynamic pricing with strategic customers. Draft available at `http://goo.gl/GqX3Pg`.

Cheng, HK, Z Li, A Naranjo. 2016. Research notecloud computing spot pricing dynamics: Latency and limits to arbitrage. *Inform. Syst. Res.* **27**(1) 145–165.

Courcoubetis, CA, A Dimakis, MI Reiman. 2001. Providing bandwidth guarantees over a best-effort network: call-admission and pricing. *Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 1. IEEE, 459–467.

Economists, The. 2009. The mother of invention: Network operators in the poor world are cutting costs and increasing access in innovative ways. `http://goo.gl/K7b8xX`.

Gallego, G, G Van Ryzin. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.* **45**(1) 24–41.

Gans, N, S Savin. 2007. Pricing and capacity rationing for rentals with uncertain durations. *Manage. Sci.* **53**(3) 390–407.

Golrezaei, N, H Nazerzadeh, P Rusmevichientong. 2014. Real-time optimization of personalized assortments. *Manag. Sci.* **60**(6) 1532–1551.

GreenBiz. 2014. Zipcar founder robin chase: Whats next for urban mobility? `https://goo.gl/HwFx1Q`.

Hampshire, RC, WA Massey, Q Wang. 2009. Dynamic pricing to control loss systems with quality of service targets. *Probab. Eng. Inform. Sc.* **23**(02) 357–383.

Hu, Z, JH Kim, J Wang, J Byrne. 2015. Review of dynamic pricing programs in the us and europe: Status quo and policy recommendations. *Renew. Sust. Energ. Rev.* **42** 743–751.

Jain, A, K Moinzadeh, A Dumrongsiri. 2015. Priority allocation in a rental model with decreasing demand. *M&SOM* **17**(2) 236–248.

Jasin, S. 2014. Reoptimization and self-adjusting price control for network revenue management. *Oper. Res.* **62**(5) 1168–1178.

Kaspi, M, T Raviv, M Tzur, H Galili. 2016. Regulating vehicle sharing systems through parking reservation policies: Analysis and performance bounds. *Eur. J. Oper. Res.* **251**(3) 969–987.

Kelly, FP. 1991. Effective bandwidths at multi-class queues. *Queueing Syst.* **9**(1-2) 5–15.

Kim, J, R Randhawa. 2015. Asymptotically optimal dynamic pricing in queueing systems. Draft available at SSRN 2546480.

Klein, Robert, Michael Neugebauer, Dimitri Ratkovitch, Claudius Steinhardt. 2015. Differentiated time slot pricing under routing considerations in attended home delivery. Draft available at SSRN 2674061.

Lanning, S, WA Massey, B Rider, W Qiong. 1999. Optimal pricing in queueing systems with quality of service constraints. *Teletraf. Sci. Eng.* 747–756.

Lei, Y, S Jasin, A Sinha. 2016. Dynamic joint pricing and order fulfillment for e-commerce retailers. Draft available at SSRN 2753810.

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.

Levi, R, C Shi. 2015. Dynamic allocation problems in loss network systems with advanced reservation. Draft available at arXiv:1505.03774.

Liu, Y, WL Cooper. 2015. Optimal dynamic pricing with patient customers. *Oper. Res.* **63**(6) 1307–1319.

Low, DW. 1974. Optimal dynamic pricing policies for an M/M/s queue. *Oper. Res.* **22**(3) 545–561.

Maglaras, C. 2006. Revenue management for a multiclass single-server queue via a fluid model analysis. *Oper. Res.* **54**(5) 914–932.

Maglaras, C, A Zeevi. 2005. Pricing and design of differentiated services: Approximate analysis and structural insights. *Oper. Res.* **53**(2) 242–262.

Miller, BL. 1969. A queueing reward system with several customer classes. *Manage. Sci.* **16**(3) 234–245.

Moallemi, CC, M Saglam. 2012. Dynamic portfolio choice with linear rebalancing rules. *Draft available at SSRN 2011605* .

O'Mahony, ED. 2015. Smarter tools for (Citi) bike sharing. Ph.D. thesis, Cornell University.

Örmeci, EL, A Burnetas, J van der Wal. 2001. Admission policies for a two class loss system. *Stoch. Models* **17**(4) 513–540.

Özer, Ö, R Phillips. 2012. *The Oxford handbook of pricing management*. Oxford University Press.

Papier, F, UW Thonemann. 2010. Capacity rationing in stochastic rental systems with advance demand information. *Oper. Res.* **58**(2) 274–288.

Paschalidis, IC, JN Tsitsiklis. 2000. Congestion-dependent pricing of network services. *IEEE ACM Trans. Network* **8**(2) 171–184.

Pfrommer, J, J Warrington, G Schildbach, M Morari. 2014. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Trans. Intelli. Transp.* **15**(4) 1567–1578.

Project44. 2015. Project44 introduces dynamic pricing technology to freight industry. `http://goo.gl/Ayk2mG`.

Raviv, T, O Kolka. 2013. Optimal inventory management of a bike-sharing station. *IIE Trans.* **45**(10) 1077–1093.

Savin, SV, MA Cohen, N Gans, Z Katalan. 2005. Capacity management in rental businesses with two customer bases. *Oper. Res.* **53**(4) 617–631.

Schuijbroek, J, R Hampshire, WJ van Hoeve. 2013. Inventory rebalancing and vehicle routing in bike sharing systems. Draft available at http://goo.gl/AtI2V8.

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

Talluri, KT, GJ van Ryzin. 2006. *The theory and practice of revenue management*, vol. 68. Springer.

Waserhole, A. 2014. Vehicle sharing systems pricing optimization. Ph.D. thesis, Université de Grenoble.

Williams, D. 1991. *Probability with martingales*. Cambridge university press.

Xu, H, B Li. 2013. Dynamic cloud pricing for revenue maximization. *IEEE T. Cloud Comp.* **1**(2) 158–171.

Yoon, S, ME Lewis. 2004. Optimal pricing and admission control in a queueing system with periodically varying parameters. *Queueing Syst.* **47**(3) 177–199.

EC.1. Proof of Lemma 1

Consider any admissible control $\pi \in \Pi$. Per our notations above, π essentially corresponds to the demand rate sequence $\{\lambda_t^\pi\}_{t=1}^T$. By definition, the sequence $\{\lambda_t^\pi\}_{t=1}^T$ satisfies the capacity constraints in DET. Moreover, we know from Assumption A4 and Jensen's inequality that

$$\mathbf{E}\left[\sum_{t=1}^T r(\lambda_t^\pi)\right] = \sum_{t=1}^T \mathbf{E}\left[p_t(\lambda_t^\pi)\cdot D_t(\lambda_t^\pi)\right] = \sum_{t=1}^T \mathbf{E}\left[\mathbf{E}\left[p_t(\lambda_t^\pi)\cdot D_t(\lambda_t^\pi)|\mathcal{H}_t\right]\right] = \sum_{t=1}^T \mathbf{E}\left[r(\lambda_t^\pi)\right] \le \sum_{t=1}^T r\left(\mathbf{E}\left[\lambda_t^\pi\right]\right).$$

Therefore, we conclude that $J^* = J^{\pi^*} \le J^D$. ■

EC.2. Proof of Theorem 3

The proof of Theorem 3 follows similar arguments as in the proofs of Theorems 1 and 2. We still proceed in two steps: In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t \ge 1$ and $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch$(\mathbf{m}, \boldsymbol{\epsilon})$.

Step 1

We start with the first step. For some $\delta_k = o(m_k)$, whose value is to be determined later, define a sequence of events $\{\mathcal{A}_{k,i}(\epsilon_k, \delta_k)\}$ as follows:

$$\mathcal{A}_{k,i}(\epsilon_k,\delta_k) \;=\; \left\{ \max_{t\le i m_k} \left| \sum_{s=(i-1)m_k+1}^{t} \Delta_{s,k} \right| \;<\; \delta_k \right\} \quad \text{for all } i=1,\ldots,\tfrac{T}{m_k}, k=1,\ldots,K. \tag{EC.1}$$

Analogous to (4), it can be shown that

$$\mathbf{P}\left(\bar{\mathcal{A}}_{k,i}(\epsilon_k,\delta_k)\right) \le 2\cdot \exp\{r_k^2\ \min\{C_k^D-\epsilon_k, m_k\} - r_k\delta_k\} \quad \text{for all } r_k\in[0,1]. \tag{EC.2}$$

Define $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta}) = \cap_{k=1}^K \cap_{i=1}^{T/m_k} \mathcal{A}_{i,k}(\epsilon_k,\delta_k)$, where $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_K)$. By the sub-additivity property of probability, we have:

$$\mathbf{P}(\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})) \ge 1 - 2T\ \sum_{k=1}^K \frac{\exp\{r_k^2\ \min\{C_k^D-\epsilon_k, m_k\} - r_k\delta_k\}}{m_k}. \tag{EC.3}$$

Now, we make some observations. First, on the set $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})$, we always have: $\left|\frac{\epsilon_k}{n_k}+\frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i}}\Delta_{s,k}\right| \le \frac{\epsilon_k}{n_k}+\frac{\delta_k}{m}$ for all i and k. This means that, as long as the parameters ϵ_k, δ_k, and m_k are chosen such that $\frac{\epsilon_k}{n_k}+\frac{\delta_k}{m_k}\le\min\{\varphi_L,\varphi_U\}$, the condition $\lambda^D_{t,k}-\frac{\epsilon_k}{n_k}-\frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)}}\Delta_{s,k}\in\Omega_{\lambda,k}$ in Step 2 part a of DPC-Batch is always satisfied for all t. For the remaining of the proof, we will therefore assume that $\frac{\epsilon_k}{n_k}+\frac{\delta_k}{m_k}\le\min\{\varphi_L,\varphi_U\}$. Now, suppose that $t\in\mathcal{T}_{k,i_k}$ and $\max\{1,t-n_k+1\}\in\mathcal{T}_{k,j_k}$, where $t\in[n_1,T]$. We can write the total resource consumption by the end of period t as follows:

$$
\begin{aligned}
&\sum_{k=1}^{K}\sum_{s=\max\{1,t-n_k+1\}}^{t} D_{s,k}(\hat{\mathbf{p}}^D_s)\\
&=\sum_{k=1}^{K}\left[\sum_{\substack{s\ge\max\{1,t-n_k+1\}\\ s\in\mathcal{T}_{k,j_k}}} D_{s,k}(\hat{\mathbf{p}}^D_s)+\sum_{j=j_k+1}^{i_k-1}\sum_{s\in\mathcal{T}_{k,j}} D_{s,k}(\hat{\mathbf{p}}^D_s)+\sum_{s\le t,\, s\in\mathcal{T}_{k,i_k}} D_{s,k}(\hat{\mathbf{p}}^D_s)\right]\\
&=\sum_{k=1}^{K}\sum_{\substack{s\ge\max\{1,t-n_k+1\}\\ s\in\mathcal{T}_{k,j_k}}}\left(\lambda^D_{s,k}-\frac{\epsilon_k}{n_k}-\frac{1}{m_k}\sum_{l\in\mathcal{T}_{k,j_k-1}}\Delta_{l,k}+\Delta_{s,k}\right)\\
&\quad+\sum_{k=1}^{K}\sum_{j=j_k+1}^{i_k-1}\sum_{s\in\mathcal{T}_{k,j}}\left(\lambda^D_{s,k}-\frac{\epsilon_k}{n_k}-\frac{1}{m_k}\sum_{l\in\mathcal{T}_{k,j-1}}\Delta_{l,k}+\Delta_{s,k}\right)\\
&\quad+\sum_{k=1}^{K}\sum_{s\le t,\, s\in\mathcal{T}_{k,i_k}}\left(\lambda^D_{s,k}-\frac{\epsilon_k}{n_k}-\frac{1}{m_k}\sum_{l\in\mathcal{T}_{k,i_k-1}}\Delta_{l,k}+\Delta_{s,k}\right)\\
&\le\sum_{k=1}^{K}\sum_{s=\max\{1,t-n_k+1\}}^{t}\lambda^D_{s,k}-\sum_{k=1}^{K}\frac{n_1}{n_k}\epsilon_k\\
&\quad-\sum_{k=1}^{K}\left(\frac{j_k\cdot m_k-\max\{1,t-n_k+1\}}{m_k}\right)\cdot\left(\sum_{s\in\mathcal{T}_{k,j_k-1}}\Delta_{s,k}\right)-\sum_{k=1}^{K}\sum_{\substack{s<\max\{1,t-n_k+1\}\\ s\in\mathcal{T}_{k,j_k}}}\Delta_{s,k}\\
&\quad+\sum_{k=1}^{K}\left(1-\frac{t-(i_k-1)m_k}{m_k}\right)\cdot\left(\sum_{s\in\mathcal{T}_{k,i_k-1}}\Delta_{s,k}\right)+\sum_{k=1}^{K}\sum_{s\le t,\, s\in\mathcal{T}_{k,i_k}}\Delta_{s,k}\\
&\le C-\sum_{k=1}^{K}\frac{n_1}{n_k}\epsilon_k+\sum_{k=1}^{K}\left|\sum_{s\in\mathcal{T}_{k,j_k-1}}\Delta_{s,k}\right|+\sum_{k=1}^{K}\left|\sum_{\substack{s<\max\{1,t-n_k+1\}\\ s\in\mathcal{T}_{k,j_k}}}\Delta_{s,k}\right|\\
&\quad+\sum_{k=1}^{K}\left|\sum_{s\in\mathcal{T}_{k,i_k-1}}\Delta_{s,k}\right|+\sum_{k=1}^{K}\left|\sum_{s\le t,s\in\mathcal{T}_{k,i_k}}\Delta_{s,k}\right|
\end{aligned}
$$

where the last inequality follows by the definition of i_k and j_k. On the set $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})$, for each k, each of the terms with $|\cdot|$ above is at most δ_k. So, we can bound:

$$\sum_{k=1}^{K}\sum_{s=\max\{1,t-n_k+1\}}^{t} D_{s,k}(\hat{\mathbf{p}}_s^D) \le C - \sum_{k=1}^{K}\left(\frac{n_1}{n_k}\epsilon_k - 4\delta_k\right) \quad \text{for all } t\in[n_1,T]. \tag{EC.4}$$

(For $t<n_1$, we can bound the total resource consumption by the end of period t with the total resource consumption by the end of period n_1. So, the above bound also holds.) Note that (EC.4) is the analogue of (13) in the proof of Theorem 2. An immediate choice of $\boldsymbol{\delta}$ that guarantees our resource will never run out on the set $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})$ is therefore $\delta_k = \frac{n_1\epsilon_k}{4n_k} - \frac{1}{4K}$. Given this and the assumption $\frac{\epsilon_k}{n_k} + \frac{\delta_k}{m_k} \le \min\{\varphi_L,\varphi_U\}$, we conclude that the following always hold on $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})$: (i) $C_t \ge 1$ and (ii) $\lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)-1}}\Delta_{s,k} \in \Omega_{k,\lambda}$ for all t. Consequently, $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t.

Step 2

We now ready bound the average regret of DPC-Batch$(\mathbf{m},\boldsymbol{\epsilon})$. Let $\{\mathbf{p}_t\}$ be the sequence of price vector under DPC-Batch$(\mathbf{m},\boldsymbol{\epsilon})$. As in Step 2 in the proof of Theorem 2, we have:

$$\mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}] \ge \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right] - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right)\cdot \mathbf{1}\{\bar{\mathcal{G}}(\boldsymbol{\epsilon},\boldsymbol{\delta})\}\right].$$

The second expectation after the last equality above can be bounded by $r^u T\cdot \mathbf{P}(\bar{\mathcal{G}}(\boldsymbol{\epsilon},\boldsymbol{\delta}))$ where $r^u = \max_t \max_{\boldsymbol{\lambda_t}\in\Omega_\lambda} r_t(\boldsymbol{\lambda}_t)$. As for the first expectation, by Taylor's expansion and Assumption MA6, we can bound:

$$\begin{aligned}
\mathbf{E}[r_t(\hat{\mathbf{p}}_t^D)] &= \mathbf{E}\left[r_t\left(\lambda_{t,1}^D - \frac{\epsilon_1}{n_1} - \frac{1}{m_1}\sum_{s\in\mathcal{T}_{1,i_1(t)-1}}\Delta_{s,1},\ \cdots\ ,\lambda_{t,K}^D - \frac{\epsilon_K}{n_K} - \frac{1}{m_K}\sum_{s\in\mathcal{T}_{K,i_K(t)-1}}\Delta_{s,K}\right)\right] \\
&\ge r_t\left(\boldsymbol{\lambda}_t^D\right) - \Psi\sum_{k=1}^{K}\frac{\epsilon_k}{n_k} - \Psi\cdot\sum_{k=1}^{K}\mathbf{E}\left[\left(\frac{\epsilon_k}{n_k} + \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)-1}}\Delta_{s,k}\right)^2\right] \\
&\ge r_t\left(\boldsymbol{\lambda}_t^D\right) - \Psi\sum_{k=1}^{K}\left(\frac{\epsilon_k}{n_k} + \frac{2\epsilon_k^2}{n_k^2} + \frac{2}{m_k}\right)
\end{aligned}$$

where the first inequality follows from Assumption MA6; the last inequality follows because $(x+y)^2 \le 2x^2+2y^2$ for all (x,y) and $\mathbf{E}\left[\left(\sum_{s\in\mathcal{T}_{k,i_k(t)-1}}\Delta_{s,k}\right)^2\right] \le m_k$. Putting the bounds together, for all $r_k\in[0,1]$, we have:

$$\frac{J_M^D - \mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}]}{T} \leq \frac{1}{T}\cdot\left[T\Psi\sum_{k=1}^{K}\left(\frac{\epsilon_k}{n_k}+\frac{2\epsilon_k^2}{n_k^2}+\frac{2}{m_k}\right)+r^uT\cdot\mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta))\right]$$
$$\leq \sum_{k=1}^{K}\left(\frac{\Psi\epsilon_k}{n_k}+\frac{2\Psi\epsilon_k^2}{n_k^2}+\frac{2\Psi}{m_k}+\frac{2r^uT}{m_k}\cdot\exp\{r_k^2\min\{C_k^D-\epsilon_k,m_k\}-r_k\delta_k\}\right).$$

Taking $r_k=\frac{\delta_k}{2\min\{C_k^D-\epsilon_k,m_k\}}$ and substituting $\delta_k=\frac{n_1\epsilon_k}{4n_k}-\frac{1}{4K}$ yields:

$$\frac{J_M^D - \mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}]}{T} \leq M_3\cdot\sum_{k=1}^{K}\left[\frac{\epsilon_k}{n_k}+\frac{1}{m_k}+\frac{T}{m_k}\cdot\exp\left\{-\frac{(Kn_1\epsilon_k-n_k)^2}{64K^2n_k^2\min\{C_k^D-\epsilon,m_k\}}\right\}\right]$$

for some $M_3>0$ independent of T, C, m_k, $\epsilon_k\in n_k\cdot\left[\frac{1}{Kn_1},\min\left\{1,\frac{1}{K}\cdot\frac{1+4Km_k\cdot\min\{\varphi_L,\varphi_U\}}{4m_k+n_1}\right\}\right]$, and $n_1\geq\frac{1}{K\min\{\varphi_L,\varphi_U\}}$. (Note: $\epsilon_k\geq\frac{n_k}{Kn_1}$ is needed to guarantee that $\delta_k=\frac{n_1\epsilon_k}{4n_k}-\frac{1}{4K}\geq 0$ and $n_1\geq\frac{1}{K\min\{\varphi_L,\varphi_U\}}$ is needed to guarantee that $\frac{1}{Kn_1}\leq\frac{1}{K}\cdot\frac{1+4Km_k\cdot\min\{\varphi_L,\varphi_U\}}{4m_k+n_1}$.) ■

EC.3. Proof of Theorem 4

In this section, we prove the two bounds presented in (17) and (18). The proofs of these bounds are similar, and follow similar arguments as in the proof of Theorems 2. In what follows, we first show (17) in two steps: In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t\geq 1$ and $\mathbf{p}_t=\hat{\mathbf{p}}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch(m,ϵ) followed by a brief discussion on a crucial observation for deriving the bound in (19). Finally, we will comment on which parts of the proof of (17) need to be modified to show (18).

Proof of (17): Step 1

For some $\delta=o(m)$ whose exact value is to be determined later, define $\{\mathcal{A}_{i,\ell}(\epsilon,\delta)\}$ as follows:

$$\mathcal{A}_{i,\ell}(\epsilon,\delta)=\left\{\max_{t\leq im}\left|\sum_{s=(i-1)m+1}^{t}\Delta_{s,\ell}\right|<\delta\right\}\quad\text{for all } i=1,\ldots,\tfrac{T}{m},\ell=0,\ldots,L. \tag{EC.5}$$

Analogous to (4), it can be shown that

$$\mathbf{P}\left(\bar{\mathcal{A}}_{i,\ell}(\epsilon,\delta)\right)\leq 2\cdot\exp\{r^2\min\{C-\epsilon,m\}-r\delta\}\quad\text{for all } r\in[0,1]. \tag{EC.6}$$

Define $\mathcal{G}(\epsilon,\delta)=\cap_{\ell=0}^{L}\cap_{i=1}^{T/m}\mathcal{A}_{i,\ell}(\epsilon,\delta)$. By the sub-additivity property of probability,

$$\mathbf{P}(\mathcal{G}(\epsilon,\delta))\geq 1-\frac{2T(L+1)}{m}\exp\{r^2\min\{C-\epsilon,m\}-r\delta\}. \tag{EC.7}$$

Note that, on the set $\mathcal{G}(\epsilon,\delta)$, we always have: $\left|\frac{\epsilon}{n(L+1)}+\frac{1}{m}\sum_{s\in\mathcal{T}_i}\Delta_{s,\ell}\right|\leq\frac{\epsilon}{n(L+1)}+\frac{\delta}{m}$ for all i and ℓ. This means that, as long as the parameters ϵ, δ, and m are chosen such that $\frac{\epsilon}{n(L+1)}+\frac{\delta}{m}\leq\min\{\varphi_L,\varphi_U\}$, the condition $\lambda_{t,\ell}^D-\frac{\epsilon}{n(L+1)}-\frac{1}{m}\sum_{s\in\mathcal{T}_i}\Delta_{s,\ell}\in\Omega_{\lambda,\ell}$ in Step 2 part a of DPC-Batch is always satisfied for all i. For the remaining of the proof, we will therefore assume that $\frac{\epsilon}{n(L+1)}+\frac{\delta}{m}\leq\min\{\varphi_L,\varphi_U\}$. Now, suppose that $t-\ell\in\mathcal{T}_{i_\ell}$ (if $t-\ell\leq 0$, then we set $i_\ell=0$) and $\max\{1,t-\ell-n+1\}\in\mathcal{T}_{j_\ell}$, where $n\leq t\leq T$. (For $t<n$, we can bound total resource consumption by the end of period t with the total resource consumption by the end of period n.) We can bound total consumption of resource by the end of period t as follows:

$$\begin{aligned}
&\sum_{\ell=0}^{L}\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell}D_{s,\ell}(\hat{\mathbf{p}}_s^D)\\
&=\sum_{\ell=0}^{L}\left[\sum_{\substack{s\geq\max\{1,t-\ell-n+1\}\\ s\in\mathcal{T}_{j_\ell}}}D_{s,\ell}(\hat{\mathbf{p}}_s^D)+\sum_{j=j_\ell+1}^{i_\ell-1}\sum_{s\in\mathcal{T}_j}D_{s,\ell}(\hat{\mathbf{p}}_s^D)+\sum_{s\leq t-\ell,\,s\in\mathcal{T}_{i_\ell}}D_{s,\ell}(\hat{\mathbf{p}}_s^D)\right]\\
&=\sum_{\ell=0}^{L}\sum_{\substack{s\geq\max\{1,t-\ell-n+1\}\\ s\in\mathcal{T}_{j_\ell}}}\left(\lambda_{s,\ell}^D-\frac{\epsilon}{n(L+1)}-\frac{1}{m}\sum_{v\in\mathcal{T}_{j_\ell-1}}\Delta_{v,\ell}+\Delta_{s,\ell}\right)\\
&\quad+\sum_{\ell=0}^{L}\sum_{j=j_\ell+1}^{i_\ell-1}\sum_{s\in\mathcal{T}_j}\left(\lambda_{s,\ell}^D-\frac{\epsilon}{n(L+1)}-\frac{1}{m}\sum_{v\in\mathcal{T}_{j-1}}\Delta_{v,\ell}+\Delta_{s,\ell}\right)\\
&\quad+\sum_{\ell=0}^{L}\sum_{s\leq t,\,s\in\mathcal{T}_{i_\ell}}\left(\lambda_{s,\ell}^D-\frac{\epsilon}{n(L+1)}-\frac{1}{m}\sum_{v\in\mathcal{T}_{i_\ell-1}}\Delta_{v,\ell}+\Delta_{s,\ell}\right)\\
&=\sum_{\ell=0}^{L}\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell}\lambda_{s,\ell}^D-\sum_{\ell=0}^{L}\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell}\frac{\epsilon}{n(L+1)}\\
&\quad-\sum_{\ell=0}^{L}\left(\frac{j_\ell\cdot m-\max\{1,t-\ell-n+1\}}{m}\right)\cdot\left(\sum_{s\in\mathcal{T}_{j_\ell-1}}\Delta_{s,\ell}\right)-\sum_{\ell=0}^{L}\sum_{\substack{s<\max\{1,t-\ell-n+1\},\\ s\in\mathcal{T}_{j_\ell}}}\Delta_{s,\ell}\\
&\quad+\sum_{\ell=0}^{L}\left(1-\frac{t-\ell-(i_\ell-1)m}{m}\right)\cdot\left(\sum_{s\in\mathcal{T}_{i_\ell-1}}\Delta_{s,\ell}\right)+\sum_{\ell=0}^{L}\sum_{s\leq t-\ell,s\in\mathcal{T}_{i_\ell}}\Delta_{s,\ell}
\end{aligned}$$

$$\leq C - \frac{\sum_{s=\max\{1,n-L\}}^{n} s}{n(L+1)} \cdot \epsilon + \sum_{\ell=0}^{L} \left| \sum_{s \in \mathcal{T}_{j_\ell}} \Delta_{s,\ell} \right| + \sum_{\ell=0}^{L} \left| \sum_{\substack{s < \max\{1,\, t-\ell-n+1\} \\ s \in \mathcal{T}_{j_\ell}}} \Delta_{s,\ell} \right| \qquad \text{(EC.8)}$$
$$+ \sum_{\ell=0}^{L} \left| \sum_{s \in \mathcal{T}_{i_\ell - 1}} \Delta_{s,\ell} \right| + \sum_{\ell=0}^{L} \left| \sum_{s \leq t,\, s \in \mathcal{T}_{i_\ell}} \Delta_{s,\ell} \right| .$$

where the inequality follows from the definition of i_ℓ, j_ℓ, and the fact that $\sum_{\ell=0}^{L} \sum_{s=\max\{1,\,t-\ell-n+1\}}^{t-\ell} 1 \geq \sum_{s=\max\{1,n-L\}}^{n} s$ for all $t \geq n$. Note that $L \leq n$ implies

$$\sum_{s=\max\{1,n-L\}}^{n} s = \sum_{s=n-L}^{n} s = \frac{n(n+1)}{2} - \frac{(n-L)(n-L-1)}{2} = \frac{(2n-L)(L+1)}{2} \geq \frac{n(L+1)}{2}.$$

Moreover, on the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$, the terms with $|\cdot|$ in (EC.8) are all bounded by δ. Thus, we have

$$\sum_{\ell=0}^{L} \sum_{s=\max\{1,\,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq C - \frac{1}{2} \cdot \epsilon + 4(L+1)\delta \quad \text{for all } t \geq n \ . \qquad \text{(EC.9)}$$

(EC.9) is the analogue of (13) in the proof of Theorem 2. An immediate choice of δ that guarantees our resource will never run out on the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ is therefore $\delta = \frac{\epsilon - 1}{8(L+1)}$. Given this and the assumption $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$, we conclude that the following always hold on $\mathcal{G}(\epsilon, \delta)$: (i) $C_t \geq 1$ and (ii) $\lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \in \Omega_{\lambda,\ell}$. Consequently, we have $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t.

Proof of (17): Step 2

We now ready to bound the average regret of DPC-Batch(m, ϵ). Let $\{\mathbf{p}_t\}$ be the price sequence under DPC-Batch(m, ϵ). By the same argument as in Step 2 of the proof of Theorem 2, we have

$$\mathbf{E}[R^{DPC-Batch(m,\epsilon)}] \geq \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right] - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right) \cdot \mathbf{1}\{\bar{\mathcal{G}}(\epsilon, \delta)\}\right].$$

The second expectation after the last equality above can be bounded by $r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\boldsymbol{\epsilon}, \boldsymbol{\delta}))$ where $r^u = \max_t \max_{\boldsymbol{\lambda_t} \in \Omega_\lambda} r_t(\boldsymbol{\lambda}_t)$. As for the first expectation, by Taylor's expansion and Assumption MA6, we can bound

$$\mathbf{E}[r_t(\hat{\mathbf{p}}_t^D)] = \mathbf{E}\left[r_t\left(\lambda_{t,1}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i_0(t)-1}} \Delta_{s,0}, \cdots, \lambda_{t,L}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i_L(t)-1}} \Delta_{s,L}\right)\right]$$

$$\geq r_t(\boldsymbol{\lambda}_t^D) - \Psi\frac{\epsilon}{n} - \Psi \cdot \sum_{\ell=0}^{L} \mathbf{E}\left[\left(\frac{\epsilon}{n(L+1)} + \frac{1}{m}\sum_{s\in\mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell}\right)^2\right]$$
$$\geq r_t(\boldsymbol{\lambda}_t^D) - \Psi\left[\frac{\epsilon}{n} + \frac{2\epsilon^2}{n^2(L+1)} + \frac{2(L+1)}{m}\right]$$

where the first inequality follows from Assumption MA6; the last inequality follows because $(x+y)^2 \leq 2x^2 + 2y^2$ for all (x,y) and $\mathbf{E}\left[\left(\sum_{s\in\mathcal{T}_{i_\ell-1}} \Delta_{s,\ell}\right)^2\right] \leq m$ for all ℓ.

Putting the bounds together, for all $r \in [0,1]$, we have:

$$\frac{J_A^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T(L+1)} \leq \frac{1}{T(L+1)} \cdot \left[T\Psi\left(\frac{\epsilon}{n} + \frac{2\epsilon^2}{n^2(L+1)} + \frac{2(L+1)}{m}\right) + r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta))\right]$$
$$\leq \frac{\Psi\epsilon}{n(L+1)} + \frac{2\Psi\epsilon^2}{n^2(L+1)^2} + \frac{2\Psi}{m} + \frac{2r^u T}{m}\ \exp\{r^2\ \min\{C-\epsilon, m\} - r\delta\}.$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon,\, m\}}$ and substituting $\delta = \frac{\epsilon-1}{8(L+1)}$ yield:

$$\frac{J_A^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T(L+1)} \leq M_4\left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m}\cdot \exp\left\{-\frac{(\epsilon-1)^2}{256(L+1)^2\min\{C-\epsilon,\, m\}}\right\}\right]$$

for some $M_4 > 0$ for all $\epsilon \in \left[1, \min\left\{n(L+1), m(L+1), n\cdot\frac{8m(L+1)\min\{\varphi_L,\varphi_U\}+1}{8m+n}\right\}\right]$, T, C, m, and $L < n$. (Note that $\epsilon \leq \min\left\{m(L+1), n\cdot\frac{8m(L+1)\min\{\varphi_L,\varphi_U\}+1}{8m+n}\right\}$ and $\delta = \frac{\epsilon-1}{8(L+1)}$ imply $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$ and $r \in (0,1)$.)

Proof of (18)

We now prove the bound for the case $L > n$. The major difference between the proof of (17) and (18) lies in the way we bound the total resource consumption in (EC.8). We first discuss why this is important in dealing with large L. On the RHS of (EC.8), the negative term after C is an upper bound for negative total buffers in DPC-Batch (i.e., the term $-\frac{\epsilon}{n(L+1)}$ in the definition of $\lambda_{t,\ell}(\hat{\mathbf{p}}_t^D)$) and the remaining four positive terms is an upper bound for total random errors. If $L > n$, the term $\frac{\sum_{s=\max\{1,n-L\}}^n s}{n(L+1)}$ in (EC.8) equals $\frac{\sum_{s=1}^n s}{n(L+1)} = \frac{n+1}{2(L+1)}$ and the bound in (EC.9) becomes

$$\sum_{\ell=0}^{L}\ \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D)\ \leq\ C - \frac{n+1}{2(L+1)}\cdot\epsilon + 4(L+1)\delta.$$

Since $\epsilon \leq n(L+1)$ (otherwise $\hat{\mathbf{p}}_t^D$ is not well-defined), the size of ϵ is at most on the order of n^2. Per our argument in Step 1 in the proof of (17), δ represents an upper bound of the total errors of m Bernoulli random variables (for some m), which means that $\delta = \Omega(1)$. But then, $4(L+1)\delta$ is $\Omega(L)$ and we cannot always guarantee $C - \frac{n+1}{2(L+1)} \cdot \epsilon + 4(L+1)\delta \leq C$ for all large $L > n^2$ (i.e., we may not be able to find a feasible $\epsilon \leq n(L+1)$ such that $-\frac{n+1}{2(L+1)} \cdot \epsilon + 4(L+1)\delta \leq 0$). This calls for a more careful analysis on the bound of total resource consumption.

Note that, assuming we never apply $\bar{p}$ up to and including period $t - \ell \geq 0$, total resource consumption of type-ℓ request by the end of period t is $\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D)$. We divide our analysis into three cases: $n \leq t \leq L+1$, $L+1 < t \leq n+L$, and $t > n+L$. (For $t < n$, we can bound total resource consumption by the end of period t with the total resource consumption by the end of period n.) When $n \leq t \leq L+1$, all type-ℓ requests with $\ell \geq t$ have not consumed any resource yet. For $0 \leq \ell < t$, following similar arguments as in (EC.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \begin{cases} \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta & \text{if } 0 \leq \ell < t-n \\ \sum_{s=1}^{t-\ell} \lambda_{s,\ell}^D - \frac{t-\ell}{n(L+1)} \cdot \epsilon + 2\delta & \text{if } t-n \leq \ell < t \end{cases}$$

When $L+1 < t \leq n+L$, all type-ℓ requests (for all $\ell \in \{0,1,\ldots,L\}$) have already consumed some of the resources. Following similar arguments as in (EC.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \begin{cases} \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta & \text{if } 0 \leq \ell < t-n \\ \sum_{s=1}^{t-\ell} \lambda_{s,\ell}^D - \frac{t-\ell}{n(L+1)} \cdot \epsilon + 2\delta & \text{if } t-n \leq \ell \leq L \end{cases}$$

At last, when $t > n+L$, following similar arguments as in (EC.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta.$$

Given all the above bounds, the total resource consumption by the end of period $t \geq n$ can be bounded as follows:

$$\sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D)$$

$$\leq \begin{cases} C \; - \; \frac{2t-n+1}{2(L+1)} \cdot \epsilon \; + \; 2 \cdot (2t-n+1) \cdot \delta & \text{if } n \leq t \leq L+1 \\ C \; - \; \frac{n(L+1) - \frac{(L-t+n)(L-t+n+1)}{2}}{n(L+1)} \cdot \epsilon \; + \; 2 \cdot (t+L-n+2) \cdot \delta & \text{if } L+1 < t \leq n+L \\ C \; - \; \epsilon \; + \; 4(L+1)\delta & \text{if } n+L < t \leq T \end{cases}$$

We claim that, if we set $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ and $\epsilon \geq 2(L+1)$, total resource consumption by the end of period $t \geq n$ is at most $C-1$. To see this, when $n \leq t \leq L+1$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields

$$\frac{2t-n+1}{2(L+1)} \cdot \epsilon \; - \; 2 \cdot (2t-n+1) \cdot \delta = 2 \cdot (2t-n+1) \left(\frac{\epsilon}{4(L+1)} - \delta \right) = \frac{2t-n+1}{n} \geq 1.$$

When $L+1 < t \leq n+L$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields

$$\begin{aligned} &\frac{n(L+1) - \frac{(L-t+n)(L-t+n+1)}{2}}{n(L+1)} \cdot \epsilon \; - \; 2 \cdot (t+L-n+2) \cdot \delta \\ &\quad = \epsilon \; - \; 4(L+1)\delta \; + \; \frac{(L-t+n)}{2n(L+1)} \left[4n(L+1)\delta \; - \; (L-t+n+1) \cdot \epsilon \right] \\ &\quad = \frac{2(L+1)}{n} \; + \; \frac{(L-t+n)}{2n(L+1)} \left[(t-L-1)\epsilon \; - \; 2(L+1) \right] \\ &\quad \geq \frac{2(L+1)}{n} + \frac{(L-t+n)(t-L-2)}{n} \; \geq \; 1 \qquad \text{(because } t > L+1, \epsilon > 2(L+1)\text{).} \end{aligned}$$

Finally, when $t > n+L$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields $\epsilon \; - \; 4(L+1)\delta = 2 \cdot \frac{L+1}{n} > 1$.

Now, plug the choice of δ into (EC.7) and substituting $r = \frac{\delta}{2\min\{C-\epsilon, m\}}$, we can bound

$$\mathbf{P}(\mathcal{G}(\epsilon, \delta)) \geq 1 - \frac{2T(L+1)}{m} \; \exp\left\{ -\frac{(\epsilon - 2(L+1)/n)^2}{64(L+1)^2 \min\{C-\epsilon, m\}} \right\}.$$

The remaining arguments are the same as in Step 2 of the proof of (18). Note that $\epsilon \in (L+1) \cdot \left[2, \min\left\{n, m, \frac{4mn\min\{\varphi_L, \varphi_U\}+2}{4m+n}\right\}\right]$ ensures $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$ and $r \in (0,1)$. ■

EC.4. Additional Details to Numerical Experiment

We first give a detailed definition of LRC-k. Similar with DPC-Batch, we slice the selling horizon into batches, each of which is of size k (except for the last one), i.e. $\mathcal{T}_i = [(i-1)\cdot k+1, \min\{i \cdot k, T\}]$, for all $i = 1, \ldots, \lceil T/k \rceil$. LRC-$k$ is defined as follows

Linear Rate Control with Batch Restarting (LRC-k)

Step 1. Solve **DET** and get λ^D.

Step 2. At the beginning of each t, if $t \in \mathcal{T}_i$, do:

a. If $C_t \geq 1$ and $\lambda_t^D - \frac{1}{m}\sum_{s\in\mathcal{T}_i, s<t} \Delta_s/[\mathcal{T}_i - (s-(i-1)\cdot k)] \in \Omega_\lambda$, set $p_t = \hat{p}_t^D$ where

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{1}{m}\sum_{s\in\mathcal{T}_i \;,\; s<t} \frac{\Delta_s}{\mathcal{T}_i - (s-(i-1)\cdot k)};$$

b. Otherwise, set $p_t = \bar{p}$.

At last, we provide the numerical results of experiment 1 in Table EC.1. We only show the results LRC-n, since, compared to the other heuristics, LRC-k is similarly worse as LRC-n for any k.

Table EC.1 Expected regret of fifferent heuristics with varying n

n	DPC-0				DPC-ϵ				
	Regret	Std	AvgReg(%)	Runtime (ms)	Regret	Std	AvgReg(%)	Runtime (ms)	Opt. ϵ
500	824	0.68	2.03	0.5	420	1.27	1.04	0.5	0.29
1000	1182	0.67	1.46	1.1	580	1.26	0.72	0.9	0.32
2000	1672	0.65	1.03	2.1	773	1.18	0.48	2.0	0.26
3000	2080	0.66	0.86	3.2	955	1.20	0.39	3.0	0.26
4000	2398	0.65	0.74	4.2	1089	1.14	0.34	4.0	0.27
5000	2708	0.67	0.67	5.1	1226	1.12	0.30	4.8	0.29
6000	2983	0.66	0.61	6.4	1342	1.17	0.28	5.7	0.31
7000	3228	0.73	0.57	7.3	1433	1.27	0.25	6.4	0.31
8000	3406	0.70	0.53	8.4	1522	1.27	0.23	7.5	0.28

n	LRC-n				DPCB-ϵ				
	Regret	Std	AvgReg(%)	Runtime (ms)	Regret	Std	AvgReg(%)	Runtime (ms)	Opt. ϵ
500	13662	8.8	33.75	19.34	390	1.27	0.96	4.4	0.17
1000	28971	15.6	35.78	39.04	461	1.26	0.57	9.5	0.17
2000	60868	25.3	37.59	75.51	542	1.18	0.33	17.5	0.16
3000	91693	31.6	37.75	119.82	660	1.20	0.27	25.7	0.17
4000	123529	36.5	38.14	157.39	751	1.14	0.23	35.4	0.15
5000	156829	44.3	38.74	197.40	816	1.12	0.20	42.8	0.13
6000	183936	42.3	37.86	236.34	874	1.17	0.18	54.2	0.15
7000	222871	53.9	39.32	277.82	879	1.27	0.16	62.3	0.14
8000	248279	49.8	38.33	328.38	919	1.27	0.14	70.1	0.13