Working Paper - August 2018

Revenue Management with Repeated Customer Interactions

Andre P. Calmon
Technology and Operations Management, INSEAD, Fontainebleau, France, andre.calmon@insead.edu

Florin D. Ciocan
Technology and Operations Management, INSEAD, Fontainebleau, France, florin.ciocan@insead.edu

Gonzalo Romero
Rotman School of Management, University of Toronto, Toronto, Canada, gonzalo.romero@rotman.utoronto.ca

Motivated by online advertising, we model and analyze a revenue management problem where a platform interacts with a set of customers over a number of periods. Unlike traditional network revenue management, which treats the interaction between platform and customers as one-shot, we consider stateful customers who can dynamically change their goodwill towards the platform depending on the quality of their past interactions. Customer goodwill further determines the amount of budget that they allocate to the platform in the future. These dynamics create a trade-off between the platform myopically maximizing short-term revenues, versus maximizing the long-term goodwill of its customers to collect higher future revenues. We identify a set of natural conditions under which myopic policies that ignore the budget dynamics are either optimal or near-optimal; such simple policies are particularly desirable since they do not require the platform to learn the parameters of each customer dynamic and only rely on data that is readily available to the platform. We also show that, if these conditions do not hold, myopic and finite look-ahead policies can perform arbitrarily poorly in this repeated setting. From an optimization perspective, this is one of a few instances where myopic policies are optimal or near-optimal for a dynamic program with non-convex dynamics.

Key words: Revenue management, Analysis of algorithms, Dynamic Programming.
History: 2$^{\text{nd}}$ Draft - August 2018

1. Introduction

In many revenue management domains, including online advertising, on-demand services, hospitality and airlines, a platform sells a limited volume of products to a portfolio of customers. The associated pricing and allocation decisions can be broadly grouped under the umbrella term network revenue management (NRM). NRM has been given significant attention in the operations literature (see Talluri and Van Ryzin 2005, Phillips 2005, for surveys of the field) and deployed successfully in a wide range of commercial pricing and yield management systems.

Much of the NRM related research assumes that there is a single, short interval in which customers interact with the platform to potentially purchase a product. Hence, the NRM problem is formulated as a one-shot resource allocation problem. While this approach is well-understood and highly tractable, the one-shot treatment of the underlying system is not consistent with the concept of customer lifetime value. The idea of customer lifetime value has come to prevail in marketing (Gupta et al. 2006, Gupta and Lehmann 2008, Reinartz and Venkatesan 2008), where the key goal is to optimize decisions with respect to the aggregate value of the customer over his/her *entire* lifetime rather than during a one-shot interaction.

The customer lifetime value is relevant to numerous revenue management applications where customer interactions occur *repeatedly* over a long period of time, where the aggregate value of a profitable customer might dwarf the revenues that can be derived from a single interaction (providing the platform is able to retain its profitable customers). Thus, a decision that is seemingly optimal over the current period could in fact decrease customer goodwill towards the platform and reduce the value of their future business with the platform; such a decision could be highly sub-optimal over the entire customer's lifetime. At the same time, new technologies enable platforms to track customer trajectories and gather data on how goodwill is affected by past allocation decisions; this has made it practically feasible to incorporate such goodwill effects into a revenue management system. Our work can thus be seen as an attempt to model and then operationalize customer lifetime value considerations.

At an operational level, customer goodwill is the lever through which customer lifetime value can be managed. The customer's goodwill drives future interactions between customer and platform in addition to being dependent on their past history of interactions. We model goodwill as a *state* that evolves dynamically over time, whereby repeated interactions between platform and customers become a sequence of distinct network revenue management problems, coupled by this dynamic state variable. We call this the *repeated network revenue management (RNRM)* problem, as opposed to classic one-shot problem.

Repeated RM in online advertising. Although our model is generic, our structural assumptions are grounded in the online advertising domain. The customers are advertisers who run campaigns via online advertising platforms. A platform can repeatedly interact with a given customer over hundreds of different campaign cycles. The corresponding RNRM problem evolves in discrete time. There are T periods over which the platform interacts with a fixed pool of heterogeneous customers indexed by j. At the start of each period, the platform has a limited but known supply of heterogeneous products (impressions), indexed by i, to be allocated to the customers. Each customer j has a valuation v_{ij} and there is a pre-specified price p_{ij} for each product i that can be allocated to customer j. At the start of a single period t, each customer decides the budget $b_{j,t}$ they are willing to spend on their period t campaign. Subsequently, the platform decides how to allocate products to customers taking into account their budget constraints. In the next period, the campaign budgets are revised and the process repeats with a renewed supply of products.

A customer's state is specified by the budget $b_{j,t}$. In online advertising, where it is relatively frictionless for customers to switch between competing advertising platforms depending on their expectations of campaign fulfillment quality, the budget they commit to an advertising campaign serves as a proxy for their goodwill towards any specific platform. There is anecdotal evidence that advertisers change their budgets from one campaign to another as a function of the quality of their past experiences with the platform (Wilkens et al. 2017). Mathematically, we assume that customers update their budgets over time according to a deterministic state update function which increases or decreases the budget for the next period depending on the quality of the impression allocation received in the last campaign period. We denote the quality of the allocation that customer j receives in time t by $q_{j,t}$, which we formally define in Section 3. Informally, the budget of customer j at time $t+1$ is some function

$$b_{j,t+1} = \phi_j(b_{j,t}, q_{j,t}).$$

Given the budget dynamics specified by $\{\phi_j\}_{1\leq j\leq m}$, the task is to find a sequence of allocations that maximizes the total platform revenues over T periods.

Quantifying customer goodwill effects on the platform's bottom-line adds considerable complexity to the underlying network revenue management problem:

- Firstly, while one-shot network revenue management is very well understood and offers several prescriptions and policies, such as Talluri and Van Ryzin (1998), that are both theoretically near-optimal and practically tractable, one would not expect these to remain sound in a multi-period setting. In particular, when customer goodwill effects are large, such myopic policies which are inherently short-term looking, have the potential to leave significant revenue on the table in the aggregate.
- At the same time, more sophisticated policies aimed at capturing these effects could be impractical to implement, or even compute. One concern is the amount of customer information that more sophisticated policies would require, such as the shape of individual customer goodwill dynamics, or customer valuation information. To obtain this would involve either designing incentive-compatible mechanisms to solicit the information from customers themselves, or a procedure to estimate them from previously collected data at the customer level. While technologies such as web browser cookies allow platforms to match the identity of a customer over many time-differentiated interactions, this approach brings its own difficulties, including the cost of tracking individual customers, and the increasing governmental scrutiny of what data is collected and how it is used. [1]

[1] An example of this are the European Union's recent General Data Protection Regulation (GDPR), governing the usage of algorithms that employ customer-level data.

In light of these challenges, the central issue we address is *how to design policies for the RNRM problem that (i) admit revenue performance guarantees, and (ii) are similar, from an implementability standpoint, to existing one-shot network revenue management policies.*

1.1. Main Contributions

For the model broadly described above, our paper makes the following contributions:

Parametrized performance guarantees for the myopic policy. By ‘myopic’ we mean a policy which ignores customer dynamics and solves each individual NRM problem separately and without regard to its impact towards future periods (in other words, it treats the problem as one-shot). As alluded to in the above, such policies are desirable in terms of tractability and simplicity.

We identify sufficient regularity conditions on the RNRM problem structure such that the myopic policy is a constant factor approximation of the optimum. Intuitively, the necessary conditions require similarity in the ‘bang-per-buck’ across all the products that a given customer desires: for the goods valued by customer j, all v_{ij}/p_{ij} bang-per-buck ratios lie within a constant range that is specific to that customer j. More precisely, there exists a constant $0 \leq \gamma < 1$ such that

$$\min_i \left\{ \frac{v_{ij}}{p_{ij}} \right\} - 1 \geq (1-\gamma) \cdot \left(\max_i \left\{ \frac{v_{ij}}{p_{ij}} \right\} - 1 \right),$$

where the minimum is taken over i such that $v_{ij} > 0$. In this case, we show that *the myopic policy garners at least a* $(1-\gamma)$ *fraction of what the optimal policy can achieve*, regardless of the horizon length, number of products and customers, etc. We emphasize that the platform does not need to know the exact ratio v_{ij}/p_{ij} for all items and customers for our result to hold. Also, through numerical experiments, we show that the performance of the myopic policy is often significantly higher than the $(1-\gamma)$ bound, although there are certain problem instances where the performance of the myopic policy is near the bound.

Sufficient conditions for the optimality of the myopic policy. We show that if $\gamma = 0$, i.e. if the bang-per-buck ratios are a constant specific to each customer, then our results imply that the myopic policy is guaranteed to be *optimal* for the RNRM problem. While this assumption restricts the generality of our model, it is consistent with contracts in the online advertising industry where advertisers treat a certain population with specific features as fully homogeneous. Moreover, these results hold for *heterogeneous* budget updates chosen from a general class of functions. For example, one customer may be very sensitive to the service quality received, while another could be entirely insensitive to it.

We emphasize that both of these results are surprising, since they imply that the platform can ignore goodwill effects and apply the same prescriptions suggested by one-shot network revenue management models with a limited or no loss of optimality.

From an optimization perspective, our results on the optimality or near-optimality of myopic contribute to the dynamic optimization literature and are, to the best of our knowledge, one of the few instances where one can prove such results for a dynamic program with non-convex dynamics. The technical insight that underlies our result is that, when our assumptions hold, the problem satisfies what one could call a "dynamic complementary slackness" property (cf. Proposition 3).

Hardness results. In contrast, we show that if we remove all the regularity assumptions on bang-per-buck ratios, then both the myopic as well as any finite look-ahead heuristics can accrue an *arbitrary* loss of performance with respect to the optimal policy in the RNRM problem. To prove this we develop a family of problem instances for which we can construct feasible policies which garner arbitrarily more revenues than myopic or look-ahead, while at the same time being lower bounds for the true optimum. These lower bound policies accrue revenues in a highly non-smooth fashion, purposefully depleting customer budgets over many periods only to set themselves up for a one-time, large increase in budgets (and accordingly, revenues) at the end of the time horizon. Our bang-per-buck regularity conditions stated above, which allow performance guarantees for the myopic policy, can be interpreted as ensuring that the per-period revenue function remains smooth enough over the course of the time horizon to invalidate this type of behavior.

Although, in the worst case, look-ahead policies can perform as poorly as myopic, their ability to account for customer dynamics allows them to outperform myopic on some problem instances, though at the cost of additional informational and computational burdens.

Characterization of the efficiency of policies that use limited customer data. A crucial advantage of the myopic policy is practicality. Its implementation only requires knowledge of the current budget configuration $\mathbf{b}_t$, but not of the budget update functions $\phi_j(\cdot,\cdot)$ nor of the customer valuation vectors $\mathbf{v}_j$. Consequently, even though the underlying customer-level dynamics can be complicated and heterogeneous, the algorithm itself is completely oblivious to their specification, reducing the platform's need to track customer-level data. Alternative policies which track customer goodwill dynamics (such as, for example, look-ahead policies) presumably require knowledge of this information, which may be difficult for the platform to acquire.

More generally, our model asks whether having more granular data on customer behavior, such as the history of customer interactions with the platform and information on how customers value these interactions, can lead to better decision-making. We identify conditions such that using policies which are completely agnostic to this finer grained data induces a limited loss from the perspective of the platform's long-term revenues. Moreover, we also show that in the absence of these conditions such policies can perform arbitrarily badly. We believe that the question of whether such simple and partially data-agnostic policies can perform well for complex systems is of broad interest to the operations management literature and that our work will inspire additional research on this topic.

2. Literature Review

Our work has links to several existing streams of literature. On the one hand, our problem is solvable by dynamic programming (DP), a method that has received intense attention in recent decades. On the other hand, our problem is a network revenue management or resource allocation problem familiar to the operations management community. Lastly, there is a burgeoning literature on customer behavior in systems where there are repeated interactions. Our literature review is therefore organized around these three streams.

Dynamic programming. Our problem is formulated as a shortest path dynamic program and could, in principle, be solved via dynamic programming techniques. Solving DPs via the standard Bellman recursion suffers from the well-known "curse of dimensionality", hence research on this topic focuses on either finding heuristics to arrive at good approximations of optimal policies, or on identifying certain structural properties that guarantee simple policies perform well.

Good heuristics for solving such problems are often be designed via approximate dynamic programming (ADP), where the general approach is to construct an approximation architecture to the value function that is amenable to efficient computation. Among work on ADP methods, we point the reader to surveys by Bertsekas (1995) and Powell (2007), as well as to some recent papers on linear-optimization-based approaches to ADP, such as De Farias and Van Roy (2003, 2004), Desai et al. (2012). A particular sub-stream in the ADP literature is the idea of a weakly-coupled dynamic program, a special class of Markov Decision Processes (MDPs) which can be viewed as a collection of "easy" sub-problems linked together by a constraint on a single state variable. Hawkins (2003) and Adelman and Mersereau (2008) provide ADP-based heuristics for such problems. Bertsimas and Mišić (2016) explore this sub-problem structure but use a fluid based heuristic. Our problem has a similar weakly-coupled structure in the sense that each period is an individual sub-problem and the various periods are linked together by the vector of customer goodwills. However, the heuristics in this literature stream, such as in Adelman and Mersereau (2008), do not yield performance guarantees or computationally tractable approaches for our specific problem.[2] Instead, we focus on simple policies which can be shown to have guaranteed performance for our particular problem structure.

An alternative approach to solving DPs is to impose certain structural assumptions that guarantee the optimality of simple policies as hinted at above. Denardo and Rothblum (1983) and Sobel (1990a,b) examine DPs with affine structure and provide conditions for the optimality of myopic policies. More recently, Ning and Sobel (2018) characterize the class of decomposable affine MDPs.

[2] For example, the Lagrangian relaxation developed in Adelman and Mersereau (2008) can be shown to not be tight unless one uses time dependent dual variables.

Similar to our model, these MDPs have continuous multidimensional endogenous states and actions. Assuming polyhedral properties of the decomposed sets of feasible actions and affine dynamics and rewards, Ning and Sobel (2018) show that decomposable affine MDPs have an affine value function and an extremal optimal policy, determined by solving a system of auxiliary equations. While our conclusions are similar in spirit, the state transition in our model evolves in non-linear, non-convex fashion. To the best of our knowledge, ours are the first such results for non-linear systems.

Revenue management and online resource allocation. The classical literature on NRM and resource allocation problems considers a setting in which a platform optimizes the allocation of a finite inventory of resources to a pool of heterogeneous buyers. However, the bulk of this work models a stateless, one-shot interaction between the platform and the buyers, while we consider dynamic customers who can change their behavior from one period to the next. Another important difference is that when the supply of products to be allocated over the period is known ahead of time (or in other words, deterministic), the one-shot NRM problem can be solved to optimality via an integer program, or even a linear program, under the assumption that customer budgets are large compared to unit prices. Thus, most of the work within this research stream has focused on the online case, when the supply of products evolves in an uncertain fashion and the decision maker must sequentially make irrevocable allocation decisions as the products arrive. In contrast, our model focuses on the deterministic case, which becomes non-trivial to solve because of the statefulness of the customers.

For network revenue management, the fluid approximations of Gallego and van Ryzin (1997) and Talluri and Van Ryzin (1998) are of particular note. In the case where products arrive as a point process with known rates, these yield tractable "bid-price" control policies based on linear programming. Reiman and Wang (2008) and Jasin and Kumar (2012) provide more refined approximations and heuristics which essentially improve the rates of Gallego and van Ryzin (1997), Talluri and Van Ryzin (1998). Several alternative supply uncertainty models which fluid approximations cannot handle have been considered. Such examples include adversarial arrivals (Karp et al. 1990, Mehta et al. 2005, Golrezaei et al. 2014), random order arrivals (Devanur and Hayes 2009, Agrawal and Devanur 2015) and non-stationary fluid arrivals (Ciocan and Farias 2012, Bateni et al. 2016).

In the broader revenue management literature, the last decade has brought forth some significant results on pricing in the presence of strategic customers. In this type of work, one considers agents who may strategize when to purchase a product in anticipation of future discounts offered by the seller, as in Aviv and Pazgal (2008), Liu (2007), Borgs et al. (2014), Besbes and Lobel (2015), Chen and Farias (2015), Lobel et al. (2015). While also an attempt to understand how platform allocation and pricing policies affect customer behavior, this line of work focuses on how customers shift a purchase temporally; we do not model such effects and instead model how the platform can alter the level of customer goodwill and, ultimately, the long term profits derived from customers. Recently, Agrawal et al. (2018) investigate mechanisms for repeated auctions when an auctioneer interacts with a buyer of limited rationality, such as a finite look-ahead buyer; while elegant, their results only apply to a single buyer and unit auctions.

Customer behavior and operations interface. The lifetime value of a customer is a key metric evaluating the impact of a customer acquisition or incentive plan. As platform-based models increasingly permeate traditional service industries, there has been an interest among both practitioners and academics to incorporate this metric into a company's day-to-day operations.

A stream of papers closely related to ours focus on customers who change their goodwill towards a platform over time through an exponentially smoothed update function that weighs current and past experiences. Aflaki and Popescu (2014) examine a stylized model of a service provider interacting with independent customers who remember the quality of past service, and whose retention probability depends on the history of said service, and establish structural properties of the optimal service rate. Recently, Kanoria et al. (2018) examine a related model where a firm chooses how to exercise two different quality service modes to minimize customer churn. Adelman and Mersereau (2013) model a supplier who must allocate a finite capacity of a single type of product to multiple customers, whose demand is modulated by past fill rates; they provide ADP-based heuristics for the supplier's allocation policy. Technically, our paper differs from these in that we consider an allocation problem where both customers and products are heterogeneous, which makes characterizing the optimal policy substantially harder; also, we consider a completely deterministic model and make no large market assumptions as in classical fluid approximations where uncertainty can help the analysis by essentially smoothing the problem. Additionally, Adelman and Mersereau (2013) consider the long-run average criterion in their objective function, while we explicitly focus on maximizing the platform's revenue in a finite horizon problem,[3] allowing us to model transient effects rather than just steady state behavior of the system. Lastly, L'Ecuyer et al. (2017) consider an sponsored search platform which optimizes search result rankings to trade-off between extracting instantaneous revenues from its users, and improving their user experience to improve future search engine traffic.

There are other interesting streams of work at the behavioral and operations interface, such as on loyalty programs and how they can be integrated into more traditional revenue management frameworks, see Chun et al. (2017), Chun and Ovchinnikov (2018), as well as work on pricing with reference effects, where customers interacting with a seller repeatedly exhibit anchoring behavior which depends on past prices in a similar manner to our own budget dynamics, such as in Popescu and Wu (2007), Nasiry and Popescu (2011), and Hu et al. (2016).

[3] Our results can be naturally extended to the infinite horizon discounted criterion.

3. Model

We consider a discrete time model which carries over T distinct periods. The system is endowed with a persistent population of m heterogeneous customers. These customers interact with the same platform at every period t. In each period, the platform has a supply of n different product types to allocate to the pool of customers. We assume this supply is replenished at the start of each period to some deterministic level s_i for product i.

Valuations and prices. Customer j has a constant valuation v_{ij} for one unit of a product of type i, and we denote by $\mathbf{v}_j$ customer j's vector of valuations for products $1 \le i \le n$. The platform allocates one unit of product type i to customer j at a time-invariant price p_{ij}; we denote by $\mathbf{p}_j$ the vector of prices customer j pays per unit of products $1 \le i \le n$.

Customer state. At period t, each customer j is endowed with a budget $b_{j,t}$ to purchase products in that period. To simplify the exposition, we assume $b_{j,t} \in [0,1]$ for each customer j and period t. We further denote by $\mathbf{b}_t$ the vector of customer budgets at period t.

Control. At period t, the platform chooses the allocation $\mathbf{x}_t \in [0, s_1]^m \times \ldots \times [0, s_n]^m$ of products to customers from the feasible set $\mathbf{X}(\mathbf{b}_t)$:

$$\sum_{i=1}^{n} p_{ij} x_{ij,t} \le b_{j,t}, \quad \forall 1 \le j \le m \tag{1}$$

$$\sum_{j=1}^{m} x_{ij,t} \le s_i, \quad \forall 1 \le i \le n. \tag{2}$$

$$x_{ij,t} \ge 0, \quad \forall 1 \le i \le n, 1 \le j \le m.$$

We define the vector $\mathbf{x}_{j,t} = (x_{1j,t}, \ldots, x_{nj,t})$ to refer to the vector of allocations of products to customer j in period t, and the feasible set $\mathbf{X}_j(b_{j,t})$ to refer to the projection of $\mathbf{X}(\mathbf{b}_t)$ to customer j's allocations. Note that we allow fractional allocations of products. In applications such as online ad allocation where the volume of products transacted is large, it could be argued that the optimality gap versus the best integral solution is small. Such an assumption allows us to simplify the problem without losing any essential insights.

Platform revenues and customer utilities. In period t and for a feasible allocation $\mathbf{x}_t$, the platform garners revenues

$$R(\mathbf{x}_t) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} x_{ij,t}.$$

We assume that customer j's utility function is linear in her allocation. Specifically,

$$U_j(b_{j,t}, \mathbf{x}_{j,t}) = \begin{cases} \sum_{i=1}^{n} (v_{ij} - p_{ij}) x_{ij,t} & \text{if } \mathbf{x}_{j,t} \in \mathbf{X}_j(b_{j,t}) \\ 0 & \text{if } \mathbf{x}_{j,t} \notin \mathbf{X}_j(b_{j,t}), \end{cases}$$

Additionally, for a given budget $b_{j,t}$, we denote the maximum possible utility that customer j could attain by

$$U_j^*(b_{j,t}) = \max_{\mathbf{y}_t \in \mathbf{X}_j(b_{j,t})} U_j(b_{j,t}, \mathbf{y}_t).$$

We emphasize that if $\mathbf{y}_t^*$ is an optimal solution to this problem, i.e. it is feasible and attains $U_j^*(b_{j,t})$, then $\sum_i p_{ij} y_{i,t}^* = \min\left(b_{j,t}, \sum_{i=1}^n p_{ij} s_i\right)$. Namely, when computing the maximum possible utility for customer j in period t, either the budget constraint becomes tight or all the supply that the customer is interested in is exhausted.

Budget state dynamics. Our model of budget dynamics formalizes the concept of customer goodwill which we introduced in Section 1. We allow budgets to evolve over time as a deterministic process $\{\mathbf{b}_t\}_{1\le t\le T}$ specified by budget update functions $\phi_j(\cdot,\cdot)$ which are potentially heterogeneous across customers. We allow for a broad range of memoryless functions $\phi_j(\cdot,\cdot)$ that depend on (i) customer j's current budget and (ii) the "service quality" provided to customer j in the current period. We assume that the latter is a function of that customer's allocation in the current period $\mathbf{x}_{j,t}$ and her available budget $b_{j,t}$. Specifically, we define the service quality provided to customer j in period t as

$$q_j(b_{j,t}, \mathbf{x}_{j,t}) \triangleq \begin{cases} \dfrac{U_j(b_{j,t}, \mathbf{x}_{j,t})}{U_j^*(b_{j,t})}, & \text{if } b_{j,t} > 0 \\ 1, & \text{if } b_{j,t} = 0. \end{cases} \tag{3}$$

Intuitively, $q_j(b_{j,t}, \mathbf{x}_{j,t})$ is the ratio of the utility customer j garnered from her allocation $\mathbf{x}_{j,t}$ and the maximum utility she could have garnered at state $b_{j,t}$, equal to $U_j^*(b_{j,t})$. Clearly, the definition implies that the range of $q_j(\cdot,\cdot)$ is the interval $[0,1]$ for any $\mathbf{x}_{j,t} \in \mathbf{X}_j(b_{j,t})$.

We define the service quality q_j to be 1 when the budget level is 0; this is a corner case that can be avoided by having $b_{j,t} \in [b_j^{\min}, b_j^{\max}]$ for each customer j and period t, where $b_j^{\min} > 0$. Although this results in a more general formulation, it also requires additional notation to project the budget back and forth between $[b_j^{\min}, b_j^{\max}]$ and $[0,1]$. For the sake of clarity, we assume that $b_{j,t} \in [0,1]$ for each customer j and period t, although our results extend to this more general case.

As a result, the domain and range of the budget update function $\phi_j\left(b_{j,t}, q_j\left(b_{j,t}, \mathbf{x}_{j,t}\right)\right)$ are $\phi_j : [0,1]^2 \to [0,1]$. To denote the updated budget state vector induced by the budget update $\phi_j\left(b_{j,t}, q_j\left(b_{j,t}, \mathbf{x}_{j,t}\right)\right)$, we use the vector notation $\boldsymbol{\phi}\left(\mathbf{b}_t, \mathbf{q}\left(\mathbf{b}_t, \mathbf{x}_t\right)\right)$. Thus, budgets over two periods are linked by the equation

$$\mathbf{b}_{t+1} = \boldsymbol{\phi}\left(\mathbf{b}_t, \mathbf{q}\left(\mathbf{b}_t, \mathbf{x}_t\right)\right).$$

We note several connections between our choice of service quality metric and the the broader literature. In supply chains, the fraction of the customer's demand that is served with the available inventory, or fill rate, is a commonly monitored metric of supplier performance. For example, Gaur and Park (2007) and Adelman and Mersereau (2013) study supplier systems where customers place orders for a homogeneous good, and use fill rates as a measure of service quality in a way that is similar to our approach. One can think of our service quality metric $q_j(b_{j,t}, \mathbf{x}_{j,t})$ as a natural analogue (in the online advertising context) to the fill rate defined in Adelman and Mersereau (2013). Specifically, $U_j^*(b_{j,t})$ can be interpreted as the (best-case) utility that the advertiser demands, whereas $U_j(b_{j,t}, \mathbf{x}_{j,t})$ can be interpreted as the utility that is actually served to the advertiser under the allocation $\mathbf{x}_t$.

A last remark is that in crafting such a model of customer goodwill one could ostensibly define a different service quality metric than the one used here. We believe ours is the right choice for a few reasons:

- First, our service quality metric satisfies the condition that the metric $q_j(\cdot)$ should be increasing in the utility derived over the set of feasible allocations that customer j can receive at their current budget level.
- Secondly, our quality metric is scale-free in the sense that, if we change the scale of $\mathbf{v}$ and $\mathbf{p}$ by multiplying by a common constant (i.e., as if we converted the values of $\mathbf{v}$ and $\mathbf{p}$ into a different currency), we obtain the same quality as before; requiring scale-freeness invalidates other candidates, such as additive metrics like $U_j(b_{j,t}, \mathbf{x}_{j,t}) - U_j^*(b_{j,t})$.

For the sake of simplicity, our service quality metric additionally lies in the $[0,1]$ interval, i.e. $q_j(0) = 0$ and $q_j(U_j^*(b_{j,t})) = 1$.

Special Budget Update Case: Exponential Smoothing. While our results require some minimal structure on the budget update function of each customer $\phi_j(\cdot,\cdot)$, made precise in Assumption 1 in Section 4, our analysis accommodates a broad class of possible update functions. A natural example is an exponentially smoothed function of the form

$$\phi_j(b_j, q_j) = \alpha_j \cdot b_j + (1 - \alpha_j) \cdot q_j, \tag{4}$$

for any $(b_j, q_j) \in [0,1]^2$ and some $\alpha_j \in [0,1]$. Note that this budget update function maintains a convex combination of the current budget level and service quality.

The exponentially smoothed update of the type (4) has been frequently considered in the literature on customer behavior as a model for customer state dynamics. For example, Gaur and Park (2007), Adelman and Mersereau (2013) and Aflaki and Popescu (2014) use exponential smoothing as an update function for customer goodwill evolution, while Nasiry and Popescu (2011) and Hu et al. (2016) use it for customer reference prices.

Dynamic programming formulation. Having stated the primitives of our model, we are now ready to formulate the platform's optimization, which is to find a sequence of allocation policies $\{\mathbf{x}_t\}_{1\le t\le T}$ that solve the following problem:

$$\begin{aligned} J_T^*(\mathbf{b}_1) = \max_{\mathbf{x}_1,\ldots,\mathbf{x}_T} & \sum_{t=1}^{T} R(\mathbf{x}_t) \\ \text{s.t. } & \mathbf{x}_t \in \mathbf{X}(\mathbf{b}_t), \quad \forall t \\ & \mathbf{b}_{t+1} = \boldsymbol{\phi}(\mathbf{b}_t, \mathbf{q}(\mathbf{b}_t, \mathbf{x}_t)), \quad \forall t. \end{aligned} \tag{5}$$

We note that the cost-to-go is indexed by the problem's remaining horizon length. Problem (5) is a deterministic dynamic program which, for any initial budget $\mathbf{b}$ and feasible allocation $\mathbf{x}$, can be solved by a Bellman recursion of the form:

$$J_\tau^*(\mathbf{b}) \triangleq \max_{\mathbf{x}\in X(\mathbf{b})} \{R(\mathbf{x}) + J_{\tau-1}^*(\boldsymbol{\phi}(\mathbf{b}, \mathbf{q}(\mathbf{b}, \mathbf{x})))\}, \quad \forall\, 1 \le \tau \le T, \tag{6}$$

with the boundary condition $J_0^*(\mathbf{b}) = 0$.

Moreover, while we present the problem in a finite horizon setting, we emphasize that our results can be extended to a discounted infinite-horizon version of the problem.

Finally, for any policy π, specifying a sequence of feasible allocations $\mathbf{x}_1^\pi, \ldots, \mathbf{x}_T^\pi$, we define

$$J_T^\pi(\mathbf{b_1}) \triangleq \sum_{t=1}^{T} R(\mathbf{x}_t^\pi).$$

In particular, we now define the natural myopic policy for problem (5).

The myopic policy. As its name suggests, the myopic policy simply maximizes the platform's short-term revenue in each period, regardless of the customers' budget dynamics between periods. Thus in each period $t \in \{1, \ldots, T\}$, and given the available budgets $\mathbf{b}_t$, the myopic policy solves the following linear program:

$$\begin{aligned} \mathrm{MY}(\mathbf{b}_t) = \max_{\mathbf{x}_t} & \; R(\mathbf{x}_t) \\ \text{s.t. } & \mathbf{x}_t \in \mathbf{X}(\mathbf{b}_t). \end{aligned} \tag{7}$$

Throughout our paper, the myopic policy is denoted by the superscript MY, i.e. $\mathbf{x}_t^{MY} \in \mathbf{X}(\mathbf{b}_t)$ and $R(\mathbf{x}_t^{MY}) = \mathrm{MY}(\mathbf{b}_t)$. The myopic policy follows the natural budget update

$$\mathbf{b}_{t+1} = \phi\left(\mathbf{b}_t, \mathbf{q}\left(\mathbf{b}_t, \mathbf{x}_t^{MY}\right)\right).$$

3.1. Model Discussion

The Bellman equation (6) emphasizes that our model captures the following trade-off: in each period the platform must find a balance between myopically maximizing its short-term revenue in that period, and providing a service quality to its customers that maximizes the revenue that can be garnered in the remaining horizon. Addressing this trade-off can be difficult due to the following model features:

1. The potential heterogeneity in the bang-per-buck $\frac{v_{ij}}{p_{ij}}$ that different items i provide for each customer j and the revenue that the platform collects for their allocation p_{ij}, can create a conflict between the allocations the customers and the platform would prefer. For example, replacing a customer's preferred item by an alternative that garners more revenue in the current period can, in principle, have dramatically different effects on customer utility and, consequently, their future budget trajectories.
2. Customers can additionally be heterogeneous in their budget update functions, and in particular in the extent of their reaction to the service quality they receive, which is a function of their utility (cf. equation (3)). Some customers may be highly sensitive to receiving low service quality and drastically reduce their budget in the next period, e.g. the exponential smoothing budget update function in equation (4) with $\alpha_j = 0$, while others may have a much more stable budget update function, e.g. the exponential smoothing budget update function in equation (4) with $\alpha_j = 1$.

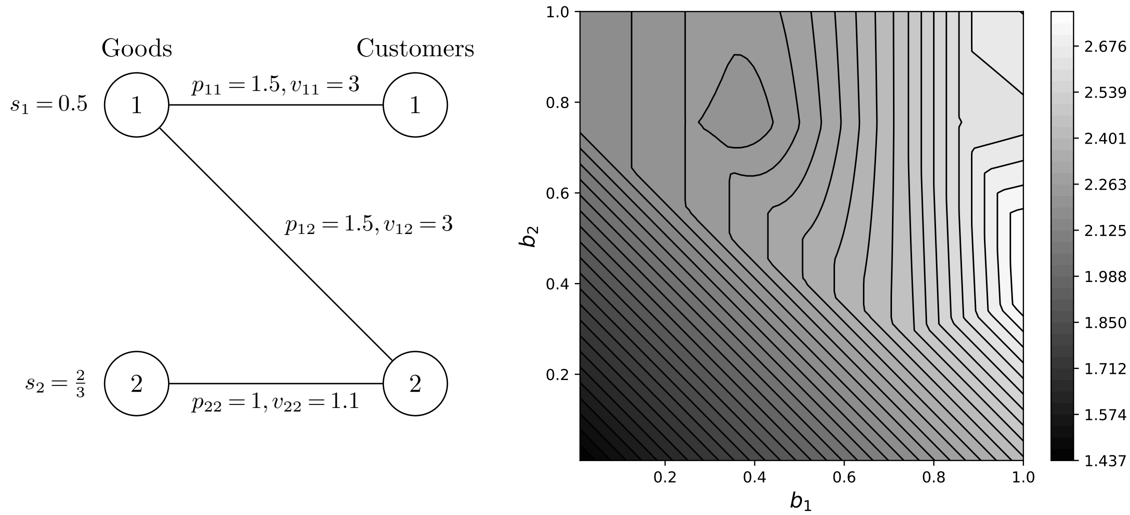
Our main results in Section 4 will address the relative importance of these challenges as they pertain to the platform's ability to maximize its revenue using relatively simple policies.

3.2. Value Function Behavior

Although our problem is a shortest path DP, it is quite challenging to solve, both from a computational, as well as a structural perspective. In this section, we provide preliminary evidence of this by exhibiting the ill behavior of the value function of problem (5), as defined in equation (6). Specifically, we show that the value function can be decreasing in customer budget levels, as well as not being quasi-concave nor quasi-convex.

First, we emphasize that in equation (6) computing $J_1(\mathbf{b}_1)$ requires solving a linear program, while computing $J_t(\mathbf{b}_t)$ for $t \in \{2, \ldots, T\}$ may require solving a non-convex optimization problem due to the constraint $\mathbf{b}_{t+1} = \boldsymbol{\phi}(\mathbf{b}_t, \mathbf{q}(\mathbf{b}_t, \mathbf{x}_t))$. In particular, this is already the case when computing $J_3(\mathbf{b}_3)$ even under the simple exponential smoothing budget update equation (4). Specifically, since

$$
\begin{aligned}
b_{j,3} &= \alpha_j b_{j,2} + (1-\alpha_j)\frac{U_j(b_{j,2}, \mathbf{x}_{j,2})}{U_j^*(b_{j,2})} \\
&= \alpha_j^2 b_{j,1} + \alpha_j(1-\alpha_j)\frac{U_j(b_{j,1}, \mathbf{x}_{j,1})}{U_j^*(b_{j,1})} + (1-\alpha_j)^2 \frac{U_j(b_{j,2}, \mathbf{x}_{j,2})}{U_j^*\left(\alpha_j b_{j,1} + (1-\alpha_j)\frac{U_j(b_{j,1}, \mathbf{x}_{j,1})}{U_j^*(b_{j,1})}\right)},
\end{aligned}
$$



(a) Model parameters. (b) Contour plot of $J_2(\mathbf{b})$.

Figure 1 Budget updates are given by equation (4)**. We assume** $\alpha_1 = \alpha_2 = 0.2$.

$J_3(\mathbf{b}_3)$ is non-convex in the first period allocation $\mathbf{x}_{j,1}$, see Lemma 1 for a numerical example. Thus, even though our problem is a shortest-path DP, it is computationally hard to find its optimum even assuming that the platform has full information about customer characteristics.

One question that arises with respect to our model is whether having customers with larger initial budgets is always beneficial for the platform. Another is whether there exist modified prices $\tilde{\mathbf{p}}$ that can incorporate the consumer's valuations into the platform's objective function such that acting myopically with respect to the modified prices $\tilde{\mathbf{p}}$ attains the optimal revenue for the platform. Specifically, it may be tempting to hope that there exist adjusted prices $\tilde{\mathbf{p}}$ which can replace the actual prices $\mathbf{p}$ in the myopic linear program, such that the policy obtained by solving

$$\max_{\mathbf{x}_t} \sum_i \sum_j \tilde{p}_{ij} x_{ij,t} \quad \text{s.t. } \mathbf{x}_t \in \mathbf{X}(\mathbf{b}_t), \tag{8}$$

at each period t is optimal. Lemma 1, stated below, answers both questions in the negative by providing an instance of problem (5) that does not satisfy either of these statements.

Lemma 1. *There exist instances of problem* (5) *such that*

(i) Increasing the customer's initial budgets decreases the platform's optimal revenue.

(ii) The revenue-to-go function is neither quasi-concave nor quasi-convex.

(iii) For any modified prices $\tilde{\mathbf{p}} \neq \mathbf{0}$*, the myopic policy with respect to* $\tilde{\mathbf{p}}$ *is guaranteed to be strictly sub-optimal.*

Proof. Consider the instance of problem (5) defined in Figure 1. It consists of two products and two customers, where the customers update their budgets according to equation (4). Moreover, Figure 1b displays the contour plots of $J_2(\mathbf{b}_2)$, which satisfy the first two statements in the lemma.

For the last statement in the lemma, it can be verified that for $T = 3$ the unique optimal allocation of products can be such that all the supply and budget constraints are non-binding. For instance, if the initial budgets are $\mathbf{b} = (0.15, 0.3)$, the unique optimal allocation in the first period is $x^*_{11,1} = 0.063$, $x^*_{12,1} = 0.01$, and $x^*_{22,1} = 0$, and it is optimal not to spend all of the customers' budgets even if there is supply available. Note that since the unique optimal solution is in the interior of the polyhedron $\mathbf{X}(\mathbf{b}_1)$, then the myopic policy will be sub-optimal for any linear objective, i.e. for any modified prices $\tilde{\mathbf{p}} \neq \mathbf{0}$. $\square$

In short, in the instance in Figure 1 it is optimal for the platform to *withhold feasible allocations*; even if supply and demand are available, it may be optimal not to match them. The intuition behind this observation comes from the temporal dynamics captured in problem (5) and the definition of the service quality in equation (3). Specifically, it may be counterproductive to let customers' budgets grow too large when the system is supply constrained, since it may then be impossible to satisfy all their demands simultaneously. Moreover, this induced low service quality may lead to disappointed customers significantly reducing their budget in future interactions with the platform. The complexity of this problem is compounded by the heterogeneity in the customers' budget update functions.

Importantly, Lemma 1(*iii*) states that, in general, the optimal policy for the RNRM problem cannot be obtained by solving any (modified) linear program at every stage as in equation (8). The unique interior-point solution from the proof of Lemma 1 cannot be the solution of an LP since it would never leave supply and demand unmatched. This shows that no linear-programming based approach can be optimal for this instance. We emphasize that the parameters of the instance in Figure 1 are reasonable for the model and do not correspond to a corner case.

4. Near-optimality of the Myopic Policy

As discussed in the introduction, the desired features of a good heuristic policy for our problem are computational simplicity and as little use of customer-specific data as possible, since this data may be difficult to acquire in practice. However, it seems unlikely that such policies would perform well in the full generality of our setting. If this were the case, it would imply that, across all instances permitted by our model, customer dynamics can be ignored no matter how drastically they can alter customer budget trajectories, suggesting that our model is degenerate. Thus, we focus on injecting realistic regularity assumptions into the model, which will allow us to show performance guarantees for simple policies.

In this section, we show that the myopic policy is, in fact, a good heuristic policy under such relatively unrestrictive regularity assumptions on the problem structure. Specifically, we derive parametric worst-case performance guarantees for the myopic policy versus the optimum of problem

(5), where the parametrization is in terms of the heterogeneity on the bang-per-buck of each item that a customer is interested in.

In our analysis, we make use of the budget trajectory induced by providing a fixed level of service quality, $q \in [0,1]$, to a customer in each period. We thus use for convenience a shorthand notation for the composition of the budget update function with itself under a fixed service quality.

Definition 1. *For each customer j and fixed service quality $q \in [0,1]$, let ϕ_j^t be defined as*

$$\phi_j^1(b,q) \triangleq \phi_j(b,q),$$
$$\phi_j^t(b,q) \triangleq \phi_j(\phi_j^{t-1}(b,q),q), \quad \forall t > 1.$$

To denote the updated budget state vector induced by the budget update $\phi_j^t(b,q)$, we use the vector notation $\boldsymbol{\phi}^t(\mathbf{b},\mathbf{q})$.

In order to derive a worst-case performance guarantee for the myopic policy, we first introduce two assumptions on the dynamics of customer budgets:

Assumption 1. *For each customer j, the budget update function $\phi_j(b,q)$ is such that:*

(i) It is non-decreasing in each component. Namely,

(a) $\phi_j(b_h,q) \geq \phi_j(b_l,q)$ for each $q \in [0,1]$ $b_h, b_l \in [0,1]$, $b_h \geq b_l$

(b) $\phi_j(b,q_h) \geq \phi_j(b,q_l)$ for each $b \in [0,1]$ $q_h, q_l \in [0,1]$, $q_h \geq q_l$.

(ii) For each $(b,q) \in [0,1]^2$, $\phi_j(b,q) \geq \min\{b,q\}$.

(iii) For each $b \in [0,1]$ and $t \in \{0,1,\ldots\}$, $\phi_j^t(b,(1-\gamma)) \geq (1-\gamma)\phi_j^t(b,1)$, for any $\gamma \in [0,1]$.

Assumption 1(i) imposes the natural condition that, for each customer j, having a larger budget or receiving a higher service quality cannot lead to a smaller budget in the next state. In particular, Assumption 1(i) implies that $\phi_j^{t-1}(b_j,1)$ is an upper bound on the budget state of customer j in period t, when starting from the initial budget b_j.

Assumption 1(ii) imposes some smoothness on the customer budget update. Specifically, the current budget can be interpreted as a summary statistic of the history of past service quality provided to the customer. Then, Assumption 1(ii) requires that the updated summary statistic cannot be lower than both its initial value and the new observation that it is being updated with. Any averaging rule will satisfy this assumption, e.g. the exponential smoothing budget update discussed in Section 3.

Finally, Assumption 1(iii) states that, for each customer, the budget induced by providing a consistent service quality of $(1-\gamma)$ is no worse than scaling by $(1-\gamma)$ the budget induced by consistently providing perfect service quality (i.e. $q = 1$). Note that the exponential smoothing budget update in equation (4) satisfies Assumption 1(iii).

Importantly, Assumption 1(i) allows us to derive the following relaxation of the platform's problem which will be useful in the proof of this section's main result: find a sequence of allocations

$\{\mathbf{x}_t\}_{1\le t\le T}$ that maximizes the total revenue collected by the platform, assuming the largest possible budget update in each period t starting from the initial budgets $\mathbf{b}_1$, i.e. assuming $b_{j,t} = \phi_j^{t-1}(b_{j,1}, 1)$ for each customer j and period $t \in \{2,\dots,T\}$. This natural relaxation can be cast in terms of the following linear program:

$$\begin{aligned} J_T^{\text{relax}}(\mathbf{b}_1) = \max_{\mathbf{x}_1,\dots,\mathbf{x}_T} & \sum_{t=1}^{T} R(\mathbf{x}_t) \\ \text{s.t. } & \mathbf{x}_t \in \mathbf{X}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1})\right), \quad \forall t. \end{aligned} \tag{9}$$

We emphasize that $J_T^{\text{relax}}(\mathbf{b}_1) = \text{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \text{MY}(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1}))$; in other words, it is an optimistic revenue upper bound which assumes all customer goodwill can be maximized at all times. The following proposition formalizes the relationship between the optimal objective value of problems (5) and (9). Its proof is presented in Appendix A.

Proposition 1. *Under Assumption 1(i), $J_T^{relax}(\mathbf{b}_1) \ge J_T^*(\mathbf{b}_1)$ for any horizon $T \ge 1$ and initial budget state $\mathbf{b}_1$.*

Note that Assumption 1 is *not sufficient* to ensure good performance of the myopic policy. In fact, the instance in the proof of Lemma 1, where the myopic policy is sub-optimal, satisfies this assumption.

We now move on to discussing our second assumption, which refers to the products' bang-per-buck ratios for each customer. This assumption yields the parametrization of our performance bound. It is motivated by the intuition that, in practice, while there may be significant heterogeneity in the goods that a given customer desires, one would expect similarity in terms of bang-per-buck between the goods in this basket. We state this precisely below:

Assumption 2. *We assume that each customer j is endowed with a characteristic set $A_j \subseteq [n]$ of products such that*

(i) $\min_{i\in A_j}\left\{\frac{v_{ij}}{p_{ij}}\right\} - 1 \ge (1-\gamma)\left(\max_{i\in A_j}\left\{\frac{v_{ij}}{p_{ij}}\right\} - 1\right)$ *for some* $\gamma \in [0,1]$,

(ii) For all $i \notin A_j, v_{ij} = p_{ij} = 0$.

Assumption 2 imposes some smoothness on the customers' preferences, which propagate to the customers' budget update through the service quality they perceive. Eventually, since the customer budgets provide an upper bound on the revenues the platform can collect, Assumption 2 rules out extremely non-smooth behavior of the platform's revenues.

Although Assumptions 1 and 2 impose structure on the general RNRM problem, a priori it is not clear this should help bound the performance of the myopic policy. Specifically,

- Even in the case that $\gamma = 0$, namely when for each customer the bang-per-buck ratio is the same across items, supply scarcity might still make it optimal to ration this supply to customers with different potentials for future revenues in a non-myopic fashion.

- Although restricting heterogeneity via Assumption 2 could, in principle, restrict the amount of sub-optimality myopic accrues over *one* period, it is not clear why this sub-optimality would not compound over time, eventually leading to a large gap versus the optimal policy.
- Moreover, the structure on the customers' budget update functions imposed by Assumption 1 still allows for high heterogeneity, which remains as an important challenge that the platform needs to address when making its allocation decisions.

The following section shows that, surprisingly, *Assumptions 1 and 2 are sufficient for the myopic policy to be* $(1-\gamma)$*-optimal in the RNRM problem.*

4.1. A Parametric Worst-Case Guarantee for the Myopic Policy

Our main result shows that simple, myopic policies admit parametric guarantees under Assumptions 1 and 2, as explained below.

Theorem 1. *For any horizon* T *and initial budget state* $\mathbf{b}_1$*, let* $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ *be the myopic policy defined in Section 3.*

Then, under Assumptions 1 and 2, $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ *is* $(1-\gamma)$*-optimal for problem* (5)*, i.e.*

$$J_T^{MY}(\mathbf{b}_1) \ge (1-\gamma) J_T^*(\mathbf{b}_1).$$

The proof of this result, which is presented in Appendix A, depends on auxiliary results which are presented in detail in the next subsection.

Theorem 1 shows that the level of heterogeneity on the bang-per-buck that each customer derives from the different products she is interested in, measured by the parameter $(1-\gamma)$ in Assumption 2, directly defines a worst-case performance guarantee for the myopic policy in maximizing the revenue collected by the platform. We emphasize that the $(1-\gamma)$ guarantee is independent of the length of the horizon T, or the number of customers m and products n. In other words, the myopic policy does not "compound" sub-optimality, with the $(1-\gamma)$ gap remaining invariant at every stage of the time horizon.

Moreover, the myopic policy has additional features that make it very attractive in practice. First, it is easy to compute since it only requires solving a linear program in each period. In fact, our result shows that continuing to use current technology built around the deterministic LP-based approaches that are common in the one-shot NRM literature is approximately optimal. In contrast, computing the optimal policy, or even a heuristic like look-ahead policies (see Section 5), may require solving a dynamic program where the revenue-to-go function is not quasi-concave nor quasi-convex, and it cannot be solved to optimality by a linear programmming based heuristic (cf. Lemma 1 and Figure 1).

Second, the myopic policy only requires knowledge of the prices, the supply of products, and the budgets available from each customer in the current period, but is agnostic to the customers' budget dynamics. In contrast, computing the optimal policy additionally requires knowledge of the customer valuations for each product and the customers' budget update functions. Although for simplicity we assumed a full-information setup where this information is available to the platform, in practice this may not be the case. In a private information setting, the platform may need to design a truthful mechanism or alternatively face the consequences of misspecified valuations.

Returning to the instance defined in Figure 1 with a horizon of $T = 3$ periods, which was our example of bad behavior of the value function in Section 3.2, we computed the actual worst-case performance of the myopic policy by full enumeration. This is attained for an initial budget of 0.01 for both customers, where the myopic policy collects about 66% of the optimal revenue. The performance of the myopic policy in this instance may suggest that the $(1-\gamma)$ bound provided in Theorem 1 is loose. We explore this issue computationally in Section 4.3.

4.2. Optimality of the Myopic Policy when $\gamma = 0$

The main purpose of this section is to present a special case of Theorem 1, namely Theorem 2 which imposes the additional condition that $\gamma = 0$, i.e. that all goods within a customer's preferred basket have constant bang-per-buck. We specifically focus on this case for two reasons (a) its additional assumptions are motivated by online advertising practice and are sufficient for establishing the optimality of myopic policies, and (b) it provides an easier to analyze "base case" for Theorem 1. Thus, in proving the optimality result for $\gamma = 0$, we will build important technical machinery that we use for proving Theorem 1, combined with some other additional, non-trivial steps.

Note that the $\gamma = 0$ assumption is practically motivated by the structure of campaign contracts in certain online advertising systems. In particular, many such systems allow advertisers to target specific features, such as geographical location, age group, or household income, that their delivered impressions must satisfy. The space of features characterizing these impression types is often quite rich, and as such it is a daunting task for an advertiser to value each particular feature combination. To avoid these complexities, campaign contracts allow an advertiser to specify a subset of the feature space, such as $(\ldots, \mathsf{age\_group} = 18-29 \text{ or } 30-39, \mathsf{state} = \mathsf{MA} \text{ or } \mathsf{NH} \text{ or } \mathsf{VT}, \ldots)$, which identifies the range of acceptable impressions for the advertiser's campaign, akin to our definition of the characteristic set A_j. It is implicitly understood that the advertiser values all the impressions in this set equally at some value $v_{ij} = v_j$, and the campaign contract sets a single price $p_{ij} = p_j$ for any impression in this set. In fact, this campaign structure is slightly more restrictive than Assumption 2 with $\gamma = 0$, since it enforces that valuations and prices are constant across one advertiser's acceptable goods basket, rather than requiring this just for their bang-per-buck ratios.

A consequence of Assumption 2 with $\gamma = 0$ is that the budget update function is simplified. Namely, denoting customer j's constant bang-per-buck by $\rho_j \triangleq \frac{v_{ij}}{p_{ij}}$ for any $i \in A_j$, the service quality experienced by customer j in period t becomes:

$$q_j(b_j, \mathbf{x}_j) = \frac{\sum_{i \in A_j}(v_{ij} - p_{ij})x_{ij}}{U_j^*(b_j)} = \frac{(\rho_j - 1)\sum_{i \in A_j} p_{ij}x_{ij}}{(\rho_j - 1)\sum_{i \in A_j} p_{ij}y_i^*} = \frac{\sum_{i \in A_j} p_{ij}x_{ij}}{\min\left(b_j, \sum_{i \in A_j} p_{ij}s_i\right)}, \tag{10}$$

where $\mathbf{y}^*$ is an optimal solution to $U_j^*(b_j)$, i.e., the service quality provided to customer j becomes the fraction of customer j's budget that was used by the platform, or her fill rate. In particular, $q_j(b_j, \mathbf{x}_j)$ becomes independent of the customer's product valuations. We also note that when $\gamma = 0$ Assumption 1(iii) becomes moot.

Having said that, we emphasize that the issues of supply scarcity and heterogeneity in customer update functions still remain. In particular, we would still expect to see the trade-off between short-term platform revenues and the long-term goodwill of the customers. For example if a customer is very sensitive to service quality, it could be better to prioritize this customer even if this contradicts the myopic policy. As a result, it is natural to expect that the myopic policy will remain sub-optimal, even in the presence of this structure that appears to simplify the problem. On the contrary, the following theorem shows that $\gamma = 0$ is sufficient for myopic optimality.

Before stating this theorem, we introduce two additional pieces of notation which we use in its proof. We use $\mu_{j,t}$ to denote an optimal dual variable associated to the budget constraint (1) of customer j, and $\lambda_{i,t}$ to denote an optimal dual variable associated to the supply constraint (2) of item i in the myopic optimization problem (7) solved at stage t.

Theorem 2. *For any horizon T and initial budget state $\mathbf{b}_1$, let $\{\mathbf{x}_t^{MY}\}_{1 \le t \le T}$ be the myopic policy defined in Section 3. Then, under Assumptions 1 and 2 with $\gamma = 0$,*

(i) $\{\mathbf{x}_t^{MY}\}_{1 \le t \le T}$ is an optimal solution to problem (9), *i.e. $J_T^{MY}(\mathbf{b}_1) = J_T^{relax}(\mathbf{b}_1)$.*

(ii) $\{\mathbf{x}_t^{MY}\}_{1 \le t \le T}$ is an optimal solution to problem (5), *i.e. $J_T^{MY}(\mathbf{b}_1) = J_T^{*}(\mathbf{b}_1)$.*

Proof. We show that $J_T^{MY}(\mathbf{b}_1) = J_T^{\text{relax}}(\mathbf{b}_1)$. Since $\{\mathbf{x}_t^{MY}\}_{1 \le t \le T}$ is feasible for problem (5), then Proposition 1 implies $J_T^{MY}(\mathbf{b}_1) = J_T^{*}(\mathbf{b}_1) = J_T^{\text{relax}}(\mathbf{b}_1)$. The proof is by induction on the horizon length T.

Base Case: If $T = 1$, then the RNRM problem collapses to the one-shot NRM problem and $J_1^{MY}(\mathbf{b}_1) = J_1^{*}(\mathbf{b}_1) = J_1^{\text{relax}}(\mathbf{b}_1)$.

Induction Step: Assume the statement of the Theorem holds for any problem with horizon length $(T-1)$, for some $T \ge 2$. Let $\{\mathbf{b}_t\}_{2 \le t \le T}$ be the budget states induced by $\{\mathbf{x}_t^{MY}\}_{1 \le t \le T-1}$. Hence, for any budget state $\mathbf{b}_2$, $\{\mathbf{x}_t^{MY}(\mathbf{b}_2)\}_{t \in \{2,\dots,T\}} \in \arg\max J_{T-1}^{*}(\mathbf{b}_2) = J_{T-1}^{\text{relax}}(\mathbf{b}_2)$, where we are abusing the notation to emphasize the dependence of the allocations $\mathbf{x}_t^{MY}$ on the budget state $\mathbf{b}_2$. In particular, since $J_{T-1}^{*}(\mathbf{b}_2) = J_{T-1}^{\text{relax}}(\mathbf{b}_2)$, it follows that $J_{T-1}^{*}(\mathbf{b}_2)$ can be computed by solving the associated linear program (9). Let $\left\{(\boldsymbol{\lambda}_t^{\text{relax}}, \boldsymbol{\mu}_t^{\text{relax}})\right\}_{2 \le t \le T}$ be associated optimal dual solutions.

For each customer j there are two possible cases:

(a) If $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} = b_{j,1}$, then it follows from Assumption 2 with $\gamma = 0$ that $q_j\left(b_{j,1}, \mathbf{x}_{j,1}^{MY}\right) = 1$ (cf. equation (10)) and $b_{j,2} = \phi_j\left(b_{j,1}, 1\right)$.

(b) If $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$, then from Proposition 3 in Appendix A it follows that $\mu_{j,2} = 0$. Moreover, from Assumption 1(i) we have $\phi^{T-2}(b_{j,2}, 1) \geq \ldots \geq b_{j,2}$. Since the optimal objective value of a maximization linear program is concave on the right hand side of the constraints, then $\mu_{j,T}^{\text{relax}} \leq \ldots \leq \mu_{j,2}^{\text{relax}} = \mu_{j,2} = 0$, see Bertsimas and Tsitsiklis (1997). From dual feasibility ($\mu_{j,t}^{\text{relax}} \geq 0$ for each $t \in \{2, \ldots, T\}$) we conclude that $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$ implies $\mu_{j,t}^{\text{relax}} = 0$ for each period $t \in \{2, \ldots, T\}$. Namely, increasing the budget $b_{j,2}$ to its upper bound $\phi_j\left(b_{j,1}, 1\right)$ (hence $\phi^{t-2}(b_{j,2}, 1)$ to $\phi^{t-1}(b_{j,1}, 1)$ for each $t \in \{2, \ldots, T\}$) does not impact the feasibility or optimality of the allocations $\{\mathbf{x}_t^{MY}(\mathbf{b}_2)\}_{2\leq t\leq T}$ in the subproblem with horizon $T-1$ starting in period $t = 2$.

Therefore, we conclude that

$$J_T^{MY}(\mathbf{b}_1) = \text{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \text{MY}\left(\boldsymbol{\phi}^{t-2}(\mathbf{b}_2, \mathbf{1})\right) = \text{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \text{MY}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1})\right) = J_T^{\text{relax}}(\mathbf{b}_1).$$

From Propositon 1, $J_T^{MY}(\mathbf{b}_1) = J_T^*(\mathbf{b}_1) = J_T^{\text{relax}}(\mathbf{b}_1)$. This concludes the proof. $\square$

Before discussing this result, we remark that a key component of the proof is our usage of Proposition 3, which informally says that, given our assumptions and $\gamma = 0$, if the myopic policy does not exhaust a customer's budget in some time period, then the marginal value of increasing that customer's budget *in the next period* is zero. In other words, Proposition 3 shows that our model satisfies what one could call "dynamic complementary slackness". This is formally proved in Appendix A.

Theorem 2(ii) implies that, under Assumptions 1 and 2, when $\gamma = 0$ an optimal policy for problem (5) with horizon T can be computed by sequentially solving T linear programs, using the myopic allocation in one period to update the customers' budget state in the next period. Moreover, Theorem 2(i) shows that under these conditions the optimal objective function of problem (5) can be computed by solving one larger linear program for the whole horizon T, where the largest possible budget update is assumed for each customer in each period, i.e. problem (9). Note that then the budget state in one period is independent of the allocation in previous periods.

Additionally, Theorem 2 shows that under Assumptions 1 and 2 some of the problematic features of the general problem discussed in Lemma 1 are no longer present when $\gamma = 0$. In particular, the platform always benefits from customers with higher budgets. This argument is formalized in the following corollary.

Corollary 1. *Under Assumptions 1 and 2 with* $\gamma = 0$, $J_T^*(\mathbf{b}_{1,h}) \geq J_T^*(\mathbf{b}_{1,l})$ *for any horizon* $T \geq 2$ *and budget states* $\mathbf{b}_{1,h}$, $\mathbf{b}_{1,l}$ *such that* $\mathbf{b}_{1,h} \geq \mathbf{b}_{1,l}$ *component-wise.*

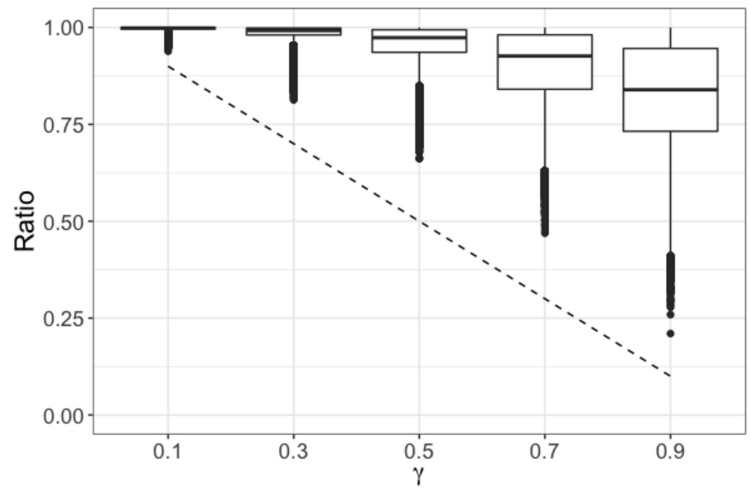


Figure 2 Boxplot of the ratio between the revenue of the myopic policy and the optimal revenue when budgets always increase at the largest possible rate. The box plot for each value of γ represents the distribution of ratios sampled from 250,000 randomly generated problem instances. The dotted line is the $1-\gamma$ bound. For these simulations, $n \in [3,10]$, $T \in [3,15]$, and $n = m$.

Proof. From Assumption 1(i) we have $\boldsymbol{\phi}^{t-1}(\mathbf{b}_{1,h}, 1) \geq \boldsymbol{\phi}^{t-1}(\mathbf{b}_{1,l}, 1)$ component-wise, for each $t \in \{2, \ldots, T\}$. Hence,

$$J_T^*(\mathbf{b}_{1,h}) = J_T^{\text{relax}}(\mathbf{b}_{1,h}) = \text{MY}(\mathbf{b}_{1,h}) + \sum_{t=2}^{T} \text{MY}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_{1,h}, \mathbf{1})\right)$$

$$\geq \text{MY}(\mathbf{b}_{1,l}) + \sum_{t=2}^{T} \text{MY}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_{1,l}, \mathbf{1})\right) = J_T^{\text{relax}}(\mathbf{b}_{1,l}) = J_T^*(\mathbf{b}_{1,l})$$

where the first and last equalities follow from Theorem 2, and the inequality follows from Lemma 2 in Appendix A. $\square$

4.3. Numerical Performance of the Myopic Policy

We now proceed to investigate the numerical performance of the myopic policy. Since computing the optimal solution to problem (5) is, in general, hard, we will use $J_T^{\text{relax}}(\mathbf{b}_1)$ as a benchmark, i.e. the revenue obtained by the platform when the budgets increase at the highest possible rate, which can be computed by solving problem (9). Recall that Proposition 1 shows that this is a valid upper bound on the optimal revenue. More specifically, assuming that customers have an exponential smoothing budget update function given by (4), we compute the performance ratio

$$r(n, m, T, \alpha, \mathbf{b}_1, \mathbf{s}, \{p_{ij}\}, \{v_{ij}\}) = \frac{J_T^{MY}(\mathbf{b}_1)}{J_T^{\text{relax}}(\mathbf{b}_1)} \tag{11}$$

for randomly generated instances of problem (5) with parameters $(n, m, T, \alpha, \mathbf{b}_1, \mathbf{s}, \{p_{ij}\}, \{v_{ij}\})$. Note that calculating $J_T^{MY}(\mathbf{b}_1)$ and $J_T^{\text{relax}}(\mathbf{b}_1)$ involves solving a sequence of linear programs.

For different values of γ, we generate 250,000 problem instances and calculate the performance ratio (11) for each instance. The parameters of each problem instance are generated as follows:

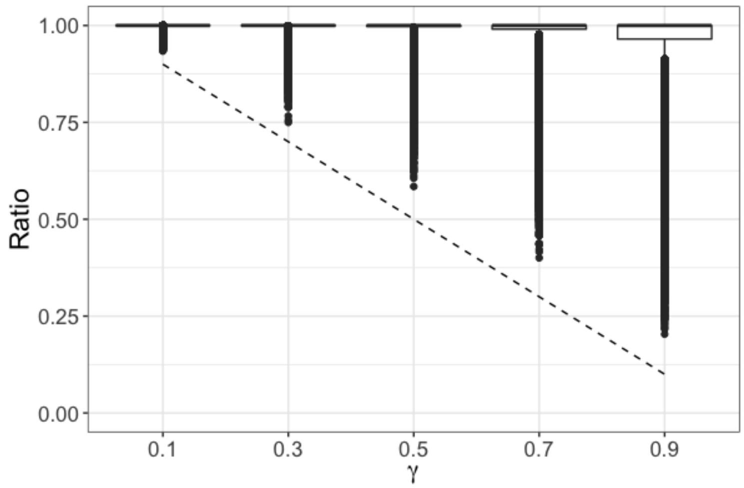


Figure 3 **Boxplot of the ratio between the revenue of the myopic policy and the optimal revenue when budgets always increase at the largest possible rate. The box plot for each value of γ represents the distribution of ratios sampled from 100,000 randomly generated problem instances. The dotted line is the $1-\gamma$ bound. For these simulations, $n=m=2$, and $T\in[3,15]$.**

- n is sampled form a discrete uniform with support $\{3,\ldots,10\}$ and we set $m=n$;
- T is sampled from a discrete uniform distribution with support $\{3,\ldots,15\}$;
- Each component of the smoothing parameters $\boldsymbol{\alpha}$, the initial budget $\mathbf{b}_1$, and the supply $\mathbf{s}$, is sampled from a uniform distribution with support $[0,1]$;
- Every customer is connected to at least one item (customer j is connected to product $i=j$). All other edges in the network are generated with some probability $\delta\in[0,1]$, where for each problem instance δ is sampled from a uniform distribution on $[0,1]$;
- For each customer j, we sample a parameter ρ from a uniform distribution with support $[0,2]$. Then the valuations $\{v_{ij}\}$ and prices and $\{p_{ij}\}$ are uniformly sampled from the circular sector defined by $(1-\gamma)\rho\leq v_{ij}/p_{ij}-1\leq\rho$ and $v_{ij}^2+p_{ij}^2\leq 25$.

The distribution of ratios r for different values of the parameter γ is depicted through box plots in Figure 2. The dotted line represents the $(1-\gamma)$ bound from Theorem 1. Note that for $\gamma\leq 0.7$, the revenue collected by the myopic policy is at least 85% of the revenue obtained when budgets increase at the highest possible rate (and therefore at least 85% of the optimal revenue) for over 75% of the randomly generated instances. However, there are many outliers in the simulation with a performance close to the $(1-\gamma)$ bound, which suggests that the bound from Theorem 1 may be tight. The majority of the outlier instances are attained for $n=3$, i.e. they have a small number of customers and products.

In order to investigate the performance of the myopic policy in small instances, we set $n=2$ and repeat the sampling procedure. The distribution of ratios for different values of γ is depicted through boxplots in Figure 3. Note that for more than 80% of the randomly generated instances the myopic policy captures over 95% of the optimal revenue. However, there are instances where the performance of the myopic policy is very close to the $(1-\gamma)$ bound, at least with respect to the revenue collected when budgets increase at the maximum possible rate.

5. Problem Instances Where Simple Policies Fail

Having determined that under certain regularity assumptions myopic policies can perform well, we now examine the performance of myopic and other natural policies if we do not impose any regularity assumption on our baseline model.

First, we define another candidate class of heuristic policies for our problem, namely finite look-ahead policies. The reason we study this class of policies is that they have been successfully used in a large number of dynamic programming applications where myopic policy performance is unsatisfactory (see for example the survey of Bertsekas 2005). Moreover, unlike myopic, finite look-ahead policies explicitly account for customer goodwill effects, and is it reasonable to expect that they would perform well. We will show in this section that, in fact, *look-ahead policies, together with the myopic policy, perform arbitrarily badly when we remove all our previous assumptions.*

We begin with the definition of a look-ahead policy.

The L-step Look-ahead Policy $\pi^{L\text{-}\mathbf{LA}}$. For any horizon T, in each period $t \leq T$ the L-step look-ahead policy $\left\{\mathbf{x}_t^{L-LA}\right\}_{1\leq t\leq T}$ implements the first period allocation of the policy that maximizes the revenue garnered by the platform in the next $\min(L+1, T-t+1)$ periods, i.e. in whatever is shorter between the current period plus the following L periods and the remaining horizon. To simplify the notation let us define $\hat{L} = \min(L+1, T-t+1)$. We note that the myopic policy defined in Section 3 is a degenerate example of an L-step look-ahead policy with $L = 0$.

Specifically, for any parameter $L \leq T-1$, period $t \leq T$, and budget state $\mathbf{b}_t$, let $\left\{\mathbf{y}_k^{\hat{L}}\right\}_{1\leq k\leq \hat{L}} \in \arg\max J_{\hat{L}}^*(\mathbf{b}_t)$. Then, $\mathbf{x}_t^{L-LA} = \mathbf{y}_1^{\hat{L}}$. We emphasize that the L-step look-ahead policy is a rolling horizon policy that requires solving a dynamic program in each period. However, it is oblivious to the customers' budget dynamics beyond the horizon $\hat{L} = \min(L+1, T-t+1)$. The L-step look-ahead policy follows the natural budget update

$$\mathbf{b}_{t+1} = \phi\left(\mathbf{b}_t, \mathbf{q}\left(\mathbf{b}_t, \mathbf{x}_t^{L-LA}\right)\right).$$

Arbitrary Sub-optimality Gap for Myopic and L-step Look-ahead Policies. In the following, we provide a family of instances of problem (5) where finite look-ahead policies, as well as myopic which is a degenerate case with $L = 0$, perform arbitrarily badly in an asymptotic regime as the number of customers of the platform grows. These results illustrate that problem (5) is hard to solve in general.

Example 1. *Consider a T period instance with $m+1$ customers. The first m customers have the same exponentially smoothed update as defined in (4), with common parameter $\alpha_j = \alpha$:*

$$\phi(b_j, q_j(b_j, \mathbf{x}_j)) = \alpha \cdot b_j + (1-\alpha) q_j(b_j, \mathbf{x}_j), \quad \forall 1 \le j \le m.$$

We set α such that $\alpha^L = 1/2$. The $(m+1)$-th customer does not update her budget, i.e. she has an exponentially smoothed update with $\alpha_{m+1} = 1$.

There are a total of $m+1$ products. The first m are indexed by $1^1, \ldots, 1^j, \ldots, 1^m$, while the last is indexed by 2. The supply of product 1^j is $s_{1^j} = \frac{1}{m^2}$, while the supply of product 2 is $s_2 = m$. We set $T = 3\log_\alpha(e^{-1})\log m$. Furthermore:

1. *For customers $1 \le j \le m$ and some scale parameter $\gamma > 2$, the product valuations are*

$$v_{1^l j} = \begin{cases} \gamma m + 1, & \text{if } l = j, \\ m + 1 - \frac{1}{m}, & \text{if } l \ne j, \end{cases} \qquad p_{1^l j} = \begin{cases} 1, & \text{if } l = j, \\ 1 - \frac{1}{m}, & \text{if } l \ne j \end{cases}$$

$$v_{2j} = 1, \quad \forall j \qquad p_{2j} = 1, \quad \forall j.$$

2. *For customer $m+1$, $v_{1^l(m+1)} = p_{1^l(m+1)} = 0$ for $1 \le l \le m$ and $v_{2(m+1)} = p_{2(m+1)} = \frac{1}{m^2}$.*
3. *For all customers $1 \le j \le m$, the starting budget is $b_{1,j} = \frac{\gamma}{\gamma+m-1}$, while for customer $(m+1)$, the starting (and constant) budget is $b_{(m+1),1} = 1$.*

Proposition 2. *On the instance described in Example 1, with $\gamma > 2$ and $m \ge \max\left\{\frac{1}{\alpha^2}, \frac{2(\gamma-1)}{\gamma-2}\right\}$, both the myopic policy and the L-step look-ahead policy, for any $L \ge 1$, produce revenues that are an $o\left(\frac{\log m}{m}\right)$ fraction of the optimum.*

The proof of proposition 2 is provided in the e-companion to this paper.

To conclude this section, we make some observations about why look-ahead policies can perform so poorly in this context. In our instance defined in Example 1, products 1^l, which are the ones with positive utility for the customers, are scarce enough that it is not feasible to simultaneously satisfy all customers $1 \le j \le m$ when their budgets are non-negligible. Thus, the feasible strategy that we exploit is to fully deplete the budgets of customers $2 \le j \le m$ to essentially 0, when it becomes possible to fully satisfy them with the scarce supply of positive utility products 1^l. This allocation induces a one-time jump in their budget from 0 to approximately $(1-\alpha)$ between periods $T-1$ and T, which is then monetized in the last period T. As a consequence, the revenues garnered by this feasible policy over the time horizon are highly non-smooth, with essentially all its revenues coming in the last period. Such a strategy cannot be found by $\pi^{L\text{-LA}}$ due to its limited look-ahead $L < T$; indeed, in only L periods it is not possible to ration the 1^l products in such a way as to cause a similar coordinated one-time jump in budgets.

6. Conclusions

In this paper, we have considered a multi-period model of network revenue management where customers' behavior changes from one period to another as a function of the quality of the past interactions between the customers and the platform controlling the allocation of products. While the problem is hard to solve in general, we have shown that by imposing reasonable conditions on the problem structure, simple myopic policies which ignore future customer behavior can work well.

We hope that our model and results inspire further efforts to understand whether classical operations management prescriptions are transferable to dynamic settings where customer interactions are repeated over time. Specifically, two avenues are worth exploring: (i) understanding how prices, along with allocations should be set in such repeated settings and (ii) building a model where customers can explicitly switch between competing platforms, depending on their history of service quality.

References

Adelman D, Mersereau AJ (2008) Relaxations of weakly coupled stochastic dynamic programs. *Operations Research* 56(3):712–727.

Adelman D, Mersereau AJ (2013) Dynamic capacity allocation to customers who remember past service. *Management Science* 59(3):592–612.

Aflaki S, Popescu I (2014) Managing retention in service relationships. *Management Science* 60(2):415–433.

Agrawal S, Daskalakis C, Mirrokni V, Sivan B (2018) Robust repeated auctions under heterogeneous buyer behavior. *arXiv preprint arXiv:1803.00494* .

Agrawal S, Devanur NR (2015) Fast algorithms for online stochastic convex programming. *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, 1405–1424.

Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management* 10(3):339–359.

Bateni MH, Chen Y, Ciocan DF, Mirrokni V (2016) Fair resource allocation in a volatile marketplace. *Proceedings of the 2016 ACM Conference on Economics and Computation*, 819–819, EC '16 (New York, NY, USA: ACM), ISBN 978-1-4503-3936-0.

Bertsekas DP (1995) *Dynamic programming and optimal control*, volume 2 (Athena Scientific Belmont, MA).

Bertsekas DP (2005) Dynamic programming and suboptimal control: A survey from adp to mpc. *European Journal of Control* 11(4-5):310–334.

Bertsimas D, Mišić VV (2016) Decomposable markov decision processes: A fluid optimization approach. *Operations Research* 64(6):1537–1555.

Bertsimas D, Tsitsiklis JN (1997) *Introduction to linear optimization* (Athena Scientific).

Besbes O, Lobel I (2015) Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* 61(1):92–110.

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

Chen Y, Farias VF (2015) Robust dynamic pricing with strategic customers. *EC*, 777.

Chun SY, Iancu D, Trichakis N (2017) Loyalty Program Liabilities and Point Values. SSRN Scholarly Paper ID 2924480, Social Science Research Network, Rochester, NY, URL `https://papers.ssrn.com/abstract=2924480`.

Chun SY, Ovchinnikov A (2018) Strategic Consumers, Revenue Management, and the Design of Loyalty Programs. SSRN Scholarly Paper ID 2606791, Social Science Research Network, Rochester, NY, URL `https://papers.ssrn.com/abstract=2606791`.

Ciocan D, Farias V (2012) Model predictive control for dynamic resource allocation. mathematics of operations research. *Mathematics of Operations Research* 37(3):501–525.

De Farias DP, Van Roy B (2003) The linear programming approach to approximate dynamic programming. *Operations research* 51(6):850–865.

De Farias DP, Van Roy B (2004) On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of operations research* 29(3):462–478.

Denardo EV, Rothblum UG (1983) Affine structure and invariant policies for dynamic programs. *Mathematics of Operations Research* 8(3):342–365.

Desai VV, Farias VF, Moallemi CC (2012) Approximate dynamic programming via a smoothed linear program. *Operations Research* 60(3):655–674.

Devanur NR, Hayes TP (2009) The adwords problem: online keyword matching with budgeted bidders under random permutations. *Proceedings of the 10th ACM conference on Electronic commerce*, 71–78 (ACM).

Gallego G, van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research* 45(1):pp. 24–41.

Gaur V, Park YH (2007) Asymmetric consumer learning and inventory competition. *Management Science* 53(2):227–240.

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

Gupta S, Hanssens D, Hardie B, Kahn W, Kumar V, Lin N, Ravishanker N, Sriram S (2006) Modeling customer lifetime value. *Journal of service research* 9(2):139–155.

Gupta S, Lehmann DR (2008) Models of customer value. *Handbook of Marketing Decision Models*, 255–290 (Springer).

Hawkins JT (2003) *A Langrangian decomposition approach to weakly coupled dynamic optimization problems and its applications*. Ph.D. thesis, Massachusetts Institute of Technology.

Hu Z, Chen X, Hu P (2016) Dynamic pricing with gain-seeking reference price effects. *Operations Research* 64(1):150–157.

Jasin S, Kumar S (2012) A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* 37(2):313–345.

Kanoria Y, Ilan L, Lu J (2018) Managing customer churn via service mode control URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3188226`.

Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, 352–358 (ACM).

Liu Y (2007) The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of Marketing* 71(4):19–35, ISSN 00222429.

Lobel I, Patel J, Vulcano G, Zhang J (2015) Optimizing product launches in the presence of strategic consumers. *Management Science* 62(6):1778–1799.

L'Ecuyer P, Maillé P, Stier-Moses NE, Tuffin B (2017) Revenue-maximizing rankings for online platforms with quality-sensitive consumers. *Operations Research* 65(2):408–423.

Mehta A, Saberi A, Vazirani U, Vazirani V (2005) Adwords and generalized on-line matching. *FOCS 2005: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 264–273 (IEEE Computer Society).

Nasiry J, Popescu I (2011) Dynamic pricing with loss-averse consumers and peak-end anchoring. *Operations research* 59(6):1361–1368.

Ning J, Sobel MJ (2018) Easy Affine Markov Decision Processes. SSRN Scholarly Paper ID 2959096, Social Science Research Network, Rochester, NY, URL `https://papers.ssrn.com/abstract=2959096`.

Phillips RL (2005) *Pricing and revenue optimization* (Stanford University Press).

Popescu I, Wu Y (2007) Dynamic pricing strategies with reference effects. *Operations Research* 55(3):413–429.

Powell WB (2007) *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703 (John Wiley & Sons).

Reiman MI, Wang Q (2008) An asymptotically optimal policy for a quantity-based network revenue management problem. *Mathematics of Operations Research* 33(2):257–282.

Reinartz WJ, Venkatesan R (2008) Decision models for customer relationship management (crm). *Handbook of marketing decision models*, 291–326 (Springer).

Sobel MJ (1990a) Higher-order and average reward myopic-affine dynamic models. *Mathematics of operations research* 15(2):299–310.

Sobel MJ (1990b) Myopic solutions of affine dynamic models. *Operations Research* 38(5):847–853.

Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management Science* 44(11):1577–1593, ISSN 0025-1909, URL `http://dx.doi.org/10.1287/mnsc.44.11.1577`.

Talluri KT, Van Ryzin GJ (2005) *The theory and practice of revenue management*. International Series in Operations Research & Management Science (New York: Springer), ISBN 0-387-24376-3.

Wilkens CA, Cavallo R, Niazadeh R (2017) Gsp: The cinderella of mechanism design. *Proceedings of the 26th International Conference on World Wide Web*, 25–32 (International World Wide Web Conferences Steering Committee).

Appendix A: Appendix for Section 4

Proposition 1. *Under Assumption 1(i), $J_T^{relax}(\mathbf{b}_1) \geq J_T^*(\mathbf{b}_1)$ for any horizon $T \geq 1$ and initial budget state $\mathbf{b}_1$.*

Let $\{\mathbf{x}_t^*\}_{t\in\{1,\ldots,T\}}$ be an optimal policy for problem (5) with initial budgets $\mathbf{b}_1$, and let $\mathbf{b}_t^*$ be the budget trajectory induced by it, i.e. $\mathbf{b}_{t+1}^* = \boldsymbol{\phi}(\mathbf{b}_t^*, \mathbf{q}(\mathbf{b}_t^*, \mathbf{x}_t^*))$ for each $t \in \{1,\ldots,T-1\}$, where $\mathbf{b}_1^* = \mathbf{b}_1$. From Assumption 1($i$), $b_{j,t}^* \leq \phi_j^{t-1}(b_{j,1}, 1)$ for each period t and customer j. Then,

$$J_T^{\text{relax}}(\mathbf{b}_1) = \text{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \text{MY}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1})\right) \geq \text{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \text{MY}(\mathbf{b}_t^*) \geq J_T^*(\mathbf{b}_1),$$

where the first inequality follows from Lemma 2 below. The second inequality follows from $\mathbf{x}_t^* \in \mathbf{X}(\mathbf{b}_t^*)$ for each period t, which implies that $R(\mathbf{x}_t^*) \leq \text{MY}(\mathbf{b}_t^*)$. □

It is not hard to see that $\text{MY}(\mathbf{b})$ is monotonically increasing in the budget state vector $\mathbf{b}$.

Lemma 2. *Let $\mathbf{b}$, $\bar{\mathbf{b}}$, be such that $\bar{b}_j \geq b_j$ for each customer j. Then, $MY(\bar{\mathbf{b}}) \geq MY(\mathbf{b})$.*

Proof. Let $\mathbf{x}^{MY}$ be a myopic policy with budgets $\mathbf{b}$. Note that $\bar{b}_j \geq b_j$ implies $\mathbf{x}^{MY} \in \mathbf{X}(\bar{\mathbf{b}})$. Hence, $\text{MY}(\bar{\mathbf{b}}) \geq R(\mathbf{x}^{MY}) = \text{MY}(\mathbf{b})$. □

The following definition will be used in the rest of the analysis.

Definition 2. *Let $\mathbf{x}_1^{MY}$ be a myopic allocation with budgets $\mathbf{b}_1$ such that there exists a customer j with a strictly loose budget constraint, i.e. such that $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$. Then, customer j defines a non-empty class of customers $C_j \subseteq \{1,\ldots,m\}$ as follows.*

Start with $C_j = \{j\}$. In each iteration define the set of customers

$$D_j(C_j) = \left\{ k \in \{1\ldots,n\} : x_{ik,1}^{MY} > 0 \text{ for some } i \in \bigcup_{l\in C_j} A_l \right\}.$$

While $D_j(C_j) \neq C_j$, update $C_j = D_j(C_j)$ and iterate.

Note that by construction the class of customers C_j has the property that all the items that the customers in C_j are interested in are allocated by $\mathbf{x}_1^{MY}$ to customers in C_j only. Namely, the class of customers C_j is such that $x_{ik,1}^{MY}=0$ for any $i\in\bigcup_{l\in C_j}A_l$, $k\notin C_j$.

Moreover, since $\sum_{i\in A_j}p_{ij}x_{ij,1}^{MY}<b_{j,1}$ it must be the case that all the items that customers in the class C_j are interested in are *fully* allocated by $\mathbf{x}_1^{MY}$ to customers in the class C_j as shown next.

Lemma 3. *Let $\mathbf{x}_1^{MY}$ be a myopic allocation with budgets $\mathbf{b}_1$ such that there exists a customer j with $\sum_{i\in A_j}p_{ij}x_{ij,1}^{MY}<b_{j,1}$, and let C_j be the associated class of customers from Definition 2, then*

$$\sum_{l\in C_j}x_{il,1}^{MY}=s_i,\ \textit{for all } i\in\bigcup_{l\in C_j}A_l.$$

Proof. Assume for a contradiction that there exists an item $i\in\bigcup_{l\in C_j}A_l$ with a strictly loose supply constraint, i.e. such that $\sum_{l\in C_j}x_{il,1}^{MY}<s_i$. We show that then there must exist an feasible augmenting path in the network induced by $\mathbf{x}_1^{MY}$, contradicting its myopic optimality.

The construction of the class C_j in Definition 2 specifies a path from customer j to item i, through items $k\in\bigcup_{l\in C_j}A_l$ and customers $l\in C_j$. Let us denote this path by $\text{path}_j(i)$. Assume, without loss of generality, that $\text{path}_j(i)$ is composed of $r\geq 1$ items labeled $\{k_1,\dots,k_r\}$, and r customers labeled $\{l_1,\dots,l_r\}$, where $l_1=j$ and $k_r=i$. Namely, $\text{path}_j(i)=\{l_1,k_1,\dots,l_r,k_r\}$. We show that $\text{path}_j(l)$ is an augmenting path.

Specifically, consider moving a flow $\gamma>0$ small enough from item i to customer j through $\text{path}_j(i)$ while keeping the spend of each customer in the path, except j, the same. Namely, denote $\gamma_r=\gamma$ and $\gamma_0=0$, and for each $s\in\{1,\dots,r\}$ consider increasing the allocation $x_{k_sl_s,1}^{MY}$ by γ_s, and decreasing the allocation $x_{k_{s-1}l_s,1}^{MY}$ by γ_{s-1}, such that $p_{k_sl_s}\gamma_s=p_{k_{s-1}l_s}\gamma_{s-1}$ for each $s\in\{2,\dots,r\}$. Equivalently, $\gamma_s=\prod_{q=s+1}^r\frac{p_{k_ql_q}}{p_{k_{q-1}l_q}}\gamma$ for each $s\in\{1,\dots,r\}$. We now verify that this change to $\mathbf{x}_1^{MY}$ is feasible for $\gamma>0$ small enough and strictly improves the revenue collected by the platform, a contradiction.

The feasibility of the change to $\mathbf{x}_1^{MY}$ is guaranteed for any $\gamma>0$ such that

$$\gamma\leq\min\left(\min_{s\in\{1,\dots,r-1\}}\left(\prod_{q=s+1}^r\frac{p_{k_{q-1}l_q}}{p_{k_ql_q}}x_{k_sl_{s+1},1}^{MY}\right),\frac{b_{j,1}-\sum_{k\in A_j}p_{kj}x_{kj,1}^{MY}}{p_{k_1j}}\prod_{q=2}^r\frac{p_{k_{q-1}l_q}}{p_{k_ql_q}},s_i-\sum_{l\in C_j}x_{il,1}^{MY}\right),$$

where the first term in the outer min guarantees the non-negativity of the modified allocations (i.e. $\gamma_s\leq x_{k_sl_{s+1},1}^{MY}$ for each $s\in\{1,\dots,r-1\}$, where $x_{k_sl_{s+1},1}^{MY}>0$ by the construction of the class C_j in Definition 2), the second term in the outer min guarantees the budget feasibility for customer j (i.e. $\sum_{i\in A_j}p_{ij}x_{ij,1}^{MY}+p_{k_1j}\gamma_1\leq b_{j,1}$), and the third term in the outer min guarantees the supply feasibility for item i.

Finally, since by construction the change to $\mathbf{x}_1^{MY}$ keeps the spend of each customer in $\text{path}_j(i)$, except j, the same, then it follows that the strict increase in the revenue collected by the platform is equal to $p_{k_1,j}\gamma_1=p_{k_1,j}\prod_{q=2}^r\frac{p_{k_ql_q}}{p_{k_{q-1}l_q}}\gamma>0$. This concludes the proof. $\square$

We are now ready to show Proposition 3, which is a key result in the proof of Theorem 2. As it will be useful for the proof which follows, we observe that the dual for problem (7), which is the linear program which can be solved to find the myopic policy, is

$$\begin{aligned}\min_{\mu_t,\lambda_t}\ &\sum_{i=1}^n s_i\lambda_{i,t}+\sum_{j=1}^m b_{j,t}\mu_{j,t}\\ \text{s.t. } &\lambda_{i,t}+p_{ij}\mu_{j,t}\geq p_{ij},\quad \forall 1\leq i\leq n,1\leq j\leq m\\ &\lambda_{i,t}\geq 0\quad \forall 1\leq i\leq n\\ &\mu_{j,t}\geq 0\quad \forall 1\leq j\leq m,\end{aligned}$$

where we remind that $\mu_{j,t}$ denotes an optimal dual variable associated to the budget constraint (1) of customer j, and $\lambda_{i,t}$ denotes an optimal dual variable associated to the supply constraint (2) of item i.

Proposition 3. *Under Assumptions 1 and 2 with $\gamma = 0$, let $\mathbf{x}_1^{MY}$ be a myopic allocation with budgets $\mathbf{b}_1$. Similarly, let $\mathbf{x}_2^{MY}$ be a myopic allocation with the updated budgets $\mathbf{b}_2 = \boldsymbol{\phi}(\mathbf{b}_1, \mathbf{q}(\mathbf{b}_1, \mathbf{x}_1^{MY}))$, and $(\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2)$ be an associated optimal dual solution.*

Then, without loss of generality $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$ implies $\mu_{j,2} = 0$, for each customer j.

Proof. Assume, for contradiction, that there exists a customer j such that $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$, *and* $\mu_{j,2} > 0$ for all dual optimal solutions associated to all myopic allocations with budgets $\mathbf{b}_2$.

First, note that $\mu_{j,2} > 0$ for *all* dual optimal solutions associated to *all* myopic allocations with budgets $\mathbf{b}_2$ implies that for a scalar $\delta > 0$ small enough, $\tilde{\mu}_{j,2} > 0$ for *all* dual optimal solutions associated to *all* myopic allocations with budgets $\tilde{\mathbf{b}}_2 = \mathbf{b}_2 + \delta \mathbf{e}_j$. In particular, let $\tilde{\mathbf{x}}_2$ be an optimal myopic allocation with budgets $\tilde{\mathbf{b}}_2$, i.e. $\tilde{\mathbf{x}}_2 \in \mathbf{X}\left(\tilde{\mathbf{b}}_2\right)$ and $R(\tilde{\mathbf{x}}_2) = \mathrm{MY}(\tilde{\mathbf{b}}_2)$.

Since $\sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < b_{j,1}$, let C_j be the associated class of customers from Definition 2. We now show that assuming $\tilde{\mu}_{j,2} > 0$ for *all* dual optimal solutions associated to *all* myopic allocations with budgets $\tilde{\mathbf{b}}_2$ implies that all the customers in the class C_j spent at least as much under allocation $\mathbf{x}_1^{MY}$ than under allocation $\tilde{\mathbf{x}}_2$, and customer j spent strictly more. Namely, we show that the class of customers C_j is such that

$$\begin{aligned} \sum_{i\in A_k} p_{ik} x_{ik,1}^{MY} \le \sum_{i\in A_k} p_{ik} \tilde{x}_{ik,2} &\text{ for each } k \in C_j, \text{ and} \\ \sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} < \sum_{i\in A_j} p_{ij} \tilde{x}_{ij,2} &\text{ for } j. \end{aligned} \tag{12}$$

This is a contradiction with $\mathbf{x}_1^{MY}$ fully allocating all the items that customers in the class C_j are interested in to customers in the class C_j (cf. Lemma 3).

To arrive at this contradiction, we *reconstruct* the class C_j as follows. We iteratively construct a class of customers $E_j \subseteq \{1, \ldots, m\}$ such that in finitely many iterations $E_j = C_j$ and (12) is satisfied. Specifically, we start with $E_j = \{j\} \subseteq C_j$, and note that $\tilde{\mu}_{j,2} > \mu_{j,1} = 0$. Then, we add a customer $l \in C_j \setminus E_j$ to the class E_j in each iteration such that the updated class preserves the properties that $\tilde{\mu}_{k,2} > \mu_{k,1} \ge 0$ for all $k \in E_j$, and $E_j \subseteq C_j$. In more details, each iteration follows the next three steps.

1. We show that since, by construction, $\tilde{\mu}_{k,2} > \mu_{k,1} \ge 0$ for all $k \in E_j$, all the customers in the class E_j spent at least as much under allocation $\mathbf{x}_1^{MY}$ than under allocation $\tilde{\mathbf{x}}_2$, and customer j spent strictly more. Namely, that the class of customers E_j is such that

$$\begin{aligned} \sum_{i\in A_k} p_{ik} x_{ik,1}^{MY} \le b_{k,2} \le \tilde{b}_{k,2} = \sum_{i\in A_k} p_{ik} \tilde{x}_{ik,2} &\text{ for each } k \in E_j, \text{ and} \\ \sum_{i\in A_j} p_{ij} x_{ij,1}^{MY} \le b_{j,2} < \tilde{b}_{j,2} = \sum_{i\in A_j} p_{ij} \tilde{x}_{ij,2} &\text{ for } j \in E_j. \end{aligned} \tag{13}$$

   In both cases in (13) the first inequality follows from Lemma 4 (where Assumptions 1 and 2 with $\gamma = 0$ are used), the second inequality follows from $\tilde{\mathbf{b}}_2 = \mathbf{b}_2 + \delta \mathbf{e}_j$ and $j \in E_j$, and the last equality follows from $\tilde{\mu}_{k,2} > \mu_{k,1} \ge 0$ for all $k \in E_j$.

2. Since. by construction, $E_j \subseteq C_j$, and from Lemma 3 all the items that customers in the class C_j are interested in are *fully* allocated by $\mathbf{x}_1^{MY}$ to customers in the class C_j, i.e. $\sum_{l \in C_j} x_{il,1}^{MY} = s_i$ for all $i \in \bigcup_{l \in C_j} A_l$, then note that (13) implies that there exists a customer $l \in C_j \setminus E_j$ and item $i \in \bigcup_{k \in E_j} A_k \cap A_l$ such that some fraction of item i that was assigned by $\mathbf{x}_1^{MY}$ to customer l is now assigned by $\tilde{\mathbf{x}}_2$ to some customer $k \in E_j$, i.e. $x_{il,1}^{MY} > \tilde{x}_{il,2} \geq 0$ and $0 \leq x_{ik,1}^{MY} < \tilde{x}_{ik,2}$ for some $l \in C_j \setminus E_j$, $k \in E_j$, and $i \in A_k \cap A_l$.
3. Consider the customers $l \in C_j \setminus E_j$ and $k \in E_j$, and the item $i \in A_k \cap A_l$, characterized in step 2. We now show that $\tilde{\mu}_{l,2} > \mu_{l,1} \geq 0$. Specifically,

$$(1-\mu_{l,1})p_{il} = \lambda_{i,1} \geq (1-\mu_{k,1})p_{ik} > (1-\tilde{\mu}_{k,2})p_{ik} = \lambda_{i,2} \geq (1-\tilde{\mu}_{l,2})p_{il}, \tag{14}$$

which implies $\tilde{\mu}_{l,2} > \mu_{l,1} \geq 0$. In (14), the first equality follows from $x_{il,1}^{MY} > 0$ and complementary slackness, the first and last inequalities follow from dual feasibility, the second inequality follows from $\tilde{\mu}_{k,2} > \mu_{k,1}$ since $k \in E_j$, the second equality follows from $\tilde{x}_{ik,2} > 0$ and complementary slackness.

Finally, update the class of customers E_j by adding customer $l \in C_j \setminus E_j$ to it, i.e. let $E_j = E_j \cup \{l\}$. Hence, the updated class E_j is such that $\tilde{\mu}_{k,2} > \mu_{k,1} \geq 0$ for all $k \in E_j$, and $E_j \subseteq C_j$. Iterate by going to step 1.

To summarize, we start with $E_j = \{j\} \subseteq C_j$ and in each iteration we add a customer $l \in C_j \setminus E_j$ to the class E_j, such that (13) is preserved. Since C_j has finitely many members, in finitely many iterations $E_j = C_j$ and (12) is satisfied. Specifically, (12) is equivalent to (13) in step 1 of the iteration for $E_j = C_j$. This is a contradiction with Lemma 3, which states that all the items that customers in the class C_j are interested in are *fully* allocated by $\mathbf{x}_1^{MY}$ to customers in the class C_j, i.e. $\sum_{l \in C_j} x_{il,1}^{MY} = s_i$ for all $i \in \bigcup_{l \in C_j} A_l$. Note that an equivalent contradiction arises in step 2 of the iteration for $E_j = C_j$, since this step implies the existence of a customer $l \in C_j \setminus C_j = \emptyset$, a contradiction. This concludes the proof. $\square$

Lemma 4. *Let $\mathbf{x}_1^{MY}$ be a myopic allocation with budgets $\mathbf{b}_1$, and let $\mathbf{b}_2$ be the updated budgets, i.e. $\mathbf{b}_2 = \boldsymbol{\phi}(\mathbf{b}_1, \mathbf{q}(\mathbf{b}_1, \mathbf{x}_1^{MY}))$. Then, under Assumptions 1 and 2 with $\gamma = 0$, $\mathbf{x}_1^{MY} \in \mathbf{X}(\mathbf{b}_2)$, i.e.*

$$\sum_{i \in A_l} p_{il} x_{il,1}^{MY} \leq b_{l,2} \text{ for each customer } l. \tag{15}$$

Proof. First, assume $b_{j,1} \leq q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY})$. Note that then

$$b_{j,2} = \phi_j(b_{j,1}, q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY})) \geq \min(b_{j,1}, q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY})) = b_{j,1} \geq \sum_{i \in A_j} p_{ij} x_{ij,1}^{MY},$$

where the first inequality follows from Assumption 1(ii).

Now assume $b_{j,1} > q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY})$. Then,

$$b_{j,2} = \phi_j(b_{j,1}, q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY})) \geq q_j(b_{j,1}, \mathbf{x}_{j,1}^{MY}) = \frac{\sum_{i \in A_j} p_{ij} x_{ij,1}^{MY}}{\min\left\{b_{j,1}, \sum_{i \in A_j} p_{ij} s_i\right\}} \geq \sum_{i \in A_j} p_{ij} x_{ij,1}^{MY}.$$

where the first inequality follows from Assumption 1(ii), the second equality follows from Assumption 2 with $\gamma = 0$ (cf. equation (10)), and the second inequality follows from $b_{j,1} \in [0,1]$. $\square$

From the proof of Theorem 2 we observe that we only require the following consequence of Assumption 2 with $\gamma = 0$: that the service quality provided to each customer satisfies equation (10). Corollary 2 shows an extension of Theorem 2 and Corollary 1 important for Theorem 1.

Corollary 2. *Under Assumption 1, for any horizon T, initial budget state $\mathbf{b}_1$, and parameter $\gamma \in [0,1)$, Theorem 2 and Corollary 1 also hold for any service quality that satisfies equation (10), and such that it is scaled in the budget update function as $\phi_j(b_j, (1-\gamma)q_j)$ for each customer j.*

Proof. It is straightforward to replicate the proofs of Theorem 2 and Corollary 1 under these assumptions. We omit the details for the sake of brevity. □

Assumption 2 allows us to derive the following lower bound for the myopic policy performance:

$$J_T^{LBMY}(\mathbf{b}_1) = \max_{\mathbf{x}_1,\ldots,\mathbf{x}_T} \sum_{t=1}^{T} R(\mathbf{x}_t) \quad \text{s.t. } \mathbf{x}_t \in \mathbf{X}\left(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1}-\boldsymbol{\gamma})\right), \quad \forall t. \tag{16}$$

Analogous to problem (9), we emphasize that $J_T^{LBMY}(\mathbf{b}_1) = \mathrm{MY}(\mathbf{b}_1) + \sum_{t=2}^{T} \mathrm{MY}(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1}-\boldsymbol{\gamma}))$.

Proposition 4. *Let $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ be the myopic policy defined in Section 3. Under Assumptions 1 and 2, $J_T^{MY}(\mathbf{b}_1) \ge J_T^{LBMY}(\mathbf{b}_1)$ for any horizon $T \ge 2$ and initial budget state $\mathbf{b}_1$.*

The proof of Proposition 4 is based on Theorem 2 and Corollary 2, as well as additional non-trival results. It is provided in the e-companion to this paper.

We are now ready to complete the proof of the main result in Section 4.

Theorem 1. *For any horizon T and initial budget state $\mathbf{b}_1$, let $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ be the myopic policy defined in Section 3.*

Then, under Assumptions 1 and 2, $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ is $(1-\gamma)$-optimal for problem (5), i.e.

$$J_T^{MY}(\mathbf{b}_1) \ge (1-\gamma) J_T^*(\mathbf{b}_1).$$

Proof. We have the following chain of inequalities,

$$J_T^{MY}(\mathbf{b}_1) \ge J_T^{LBMY}(\mathbf{b}_1) \ge (1-\gamma) J_T^{\text{relax}}(\mathbf{b}_1) \ge (1-\gamma) J_T^*(\mathbf{b}_1),$$

where the first and last inequalities follow from Propositions 4 and 1, respectively. The second inequality follows from Assumption 1(iii), since then any solution to problem (9) scaled by $(1-\gamma)$ is feasible in problem (16). Specifically, $(1-\gamma)\mathbf{X}(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1})) \subseteq \mathbf{X}(\boldsymbol{\phi}^{t-1}(\mathbf{b}_1, \mathbf{1}-\boldsymbol{\gamma}))$ for each $t \in \{1,\ldots,T\}$. This completes the proof. □

Additional Proofs

Appendix EC.1: Proofs from Section 4

Proposition 4. *Let $\{\mathbf{x}_t^{MY}\}_{1\le t\le T}$ be the myopic policy defined in Section 3. Under Assumptions 1 and 2, $J_T^{MY}(\mathbf{b}_1) \ge J_T^{LBMY}(\mathbf{b}_1)$ for any horizon $T \ge 2$ and initial budget state $\mathbf{b}_1$.*

Proof. The proof structure is as follows. We consider the revenues collected by the myopic policy under different budget updates, parametrized by an index $k \in \{2,\ldots,T\}$. Specifically, we assume that the true budget update $\mathbf{b}_{t+1} = \boldsymbol{\phi}(\mathbf{b}_t, \mathbf{q}(\mathbf{b}_t, \mathbf{x}_t^{MY}))$ is followed between the first k periods, while a modified budget update is followed in the rest of the horizon. The modified budget update is specified below, but informally it is characterized by an alternative service quality measure $(1-\gamma)\tilde{\mathbf{q}}(\mathbf{b}_t, \mathbf{x}_t^{MY})$, where $\tilde{\mathbf{q}}(\mathbf{b}_t, \mathbf{x}_t^{MY})$ satisfies equation (10). One interpretation of this alternative service quality measure is that it assumes a constant bang per buck ratio for each customer and then it scales the result by $(1-\gamma)$.

Although the net effect of using the modified budget update on the total revenue collected by the myopic policy is a priori unclear, we show that it is actually nondecreasing in the index k (cf. equation (EC.1)), i.e. the larger the number of periods where the modified budget update is used the smaller the total revenue collected by the myopic policy. This result is useful since it implies the statement in the proposition as a special case (cf. equation (EC.2)).

More precisely, let $\{\mathbf{x}_t^{MY}\}_{t\in\{1,\ldots,k\}}$ be the myopic policy starting from the initial budget state $\mathbf{b}_1$ and following the true budget update $\mathbf{b}_{t+1} = \boldsymbol{\phi}(\mathbf{b}_t, \mathbf{q}(\mathbf{b}_t, \mathbf{x}_t^{MY}))$ for each period $t \in \{1,\ldots,k-1\}$. Additionally, let $\left\{\tilde{\mathbf{x}}_{k,t}^{MY}\right\}_{t\in\{k+1,\ldots,T\}}$ be the myopic policy starting from the budget state $\mathbf{b}_k$ and following the modified budget update $\tilde{\mathbf{b}}_{k+1} = \boldsymbol{\phi}(\mathbf{b}_k, (1-\gamma)\tilde{\mathbf{q}}(\mathbf{b}_k, \mathbf{x}_k^{MY}))$, and $\tilde{\mathbf{b}}_{t+1} = \boldsymbol{\phi}\left(\tilde{\mathbf{b}}_t, (1-\gamma)\tilde{\mathbf{q}}\left(\tilde{\mathbf{b}}_t, \tilde{\mathbf{x}}_{k,t}^{MY}\right)\right)$ for each period $t \in \{k+2,\ldots,T-1\}$, where the alternative service quality function $\tilde{q}_j\left(\tilde{b}_{j,t}, \tilde{\mathbf{x}}_{j,k,t}^{MY}\right) \triangleq \frac{\sum_{i\in A_j} p_{ij}\tilde{x}_{ij,k,t}^{MY}}{\min\left(\tilde{b}_{j,t}, \sum_{i\in A_j} p_{ij}s_i\right)}$ satisfies equation (10), and it is scaled by $(1-\gamma)$.

Then, we show that the total revenue collected by the myopic policy $\left\{\{\mathbf{x}_t^{MY}\}_{t\in\{1,\ldots,k\}}, \left\{\tilde{\mathbf{x}}_{k,t}^{MY}\right\}_{t\in\{k+1,\ldots,T\}}\right\}$ is nondecreasing in the index k. Specifically, for each $k \in \{2,\ldots,T\}$,

$$\sum_{t=1}^{k} R\left(\mathbf{x}_t^{MY}\right) + \sum_{t=k+1}^{T} R\left(\tilde{\mathbf{x}}_{k,t}^{MY}\right) \ge \sum_{t=1}^{k-1} R\left(\mathbf{x}_t^{MY}\right) + \sum_{t=k}^{T} R\left(\tilde{\mathbf{x}}_{k-1,t}^{MY}\right). \tag{EC.1}$$

Therefore, in particular,

$$J_T^{MY}(\mathbf{b}_1) = \sum_{t=1}^{T} R\left(\mathbf{x}_t^{MY}\right) \ge R\left(\mathbf{x}_1^{MY}\right) + \sum_{t=2}^{T} R\left(\tilde{\mathbf{x}}_{1,t}^{MY}\right) = J_T^{LBMY}(\mathbf{b}_1), \tag{EC.2}$$

where the last equality follows from Theorem 2 and Corollary 2.

We prove equation (EC.1). The proof is by induction on the length of the horizon T.

Base Case: Consider $T=2$. Namely, we prove equation (EC.1) for $k=2$, i.e. $R\left(\mathbf{x}_1^{MY}\right)+R\left(\mathbf{x}_2^{MY}\right)\geq R\left(\mathbf{x}_1^{MY}\right)+R\left(\tilde{\mathbf{x}}_{2,2}^{MY}\right)$. We show $\tilde{\mathbf{x}}_{2,2}^{MY}\in\mathbf{X}\left(\mathbf{b}_2\right)$, hence $R\left(\tilde{\mathbf{x}}_{2,2}^{MY}\right)\leq \mathrm{MY}\left(\mathbf{b}_2\right)=R\left(\mathbf{x}_2^{MY}\right)$. Specifically,

$$\begin{aligned}
b_{j,2} &= \phi_j\left(b_{j,1}, q_j\left(b_{j,1},\mathbf{x}_{j,1}^{MY}\right)\right)\\
&= \phi_j\left(b_{j,1}, \frac{\sum_{i\in A_j}(v_{ij}-p_{ij})x_{ij,1}^{MY}}{\sum_{i\in A_j}(v_{ij}-p_{ij})y_i^*\left(b_{j,1}\right)}\right)\\
&\geq \phi_j\left(b_{j,1}, \frac{\left(\min_{i\in A_j}\left\{\frac{v_{ij}}{p_{ij}}\right\}-1\right)\sum_{i\in A_j}p_{ij}x_{ij,1}^{MY}}{\left(\max_{i\in A_j}\left\{\frac{v_{ij}}{p_{ij}}\right\}-1\right)\sum_{i\in A_j}p_{ij}y_i^*\left(b_{j,1}\right)}\right)\\
&\geq \phi_j\left(b_{j,1}, (1-\gamma)\frac{\sum_{i\in A_j}p_{ij}x_{ij,1}^{MY}}{\min\left(b_{j,1},\sum_{i\in A_j}p_{ij}s_i\right)}\right)\\
&= \phi_j\left(b_{j,1},(1-\gamma)\tilde{q}_j\left(b_{j,1},\mathbf{x}_{j,1}^{MY}\right)\right)\\
&= \tilde{b}_j^2\\
&\geq \sum_{i\in A_j}p_{ij}\tilde{x}_{ij,2,2}^{MY},
\end{aligned}$$

where the first inequality follows from Assumption 1(i), the second inequality follows from Assumptions 1(i) and 2(i), and the third inequality follows from $\tilde{\mathbf{x}}_{2,2}^{MY}\in\mathbf{X}\left(\tilde{\mathbf{b}}_2\right)$.

Induction Step: For any $T\geq 3$ assume that equation (EC.1) holds for any problem with horizon $T-1$, and for each policy defined by a parameter $k\in\{2,\ldots,T-1\}$.

First, note that this implies that equation (EC.1) holds for any problem with horizon T, and for each policy defined by a parameter $k\in\{3,\ldots,T\}$. Specifically, equation (EC.1) for a problem with horizon $(T-1)\geq 2$ and a policy defined by a parameter $k\in\{2,\ldots,T-1\}$ implies equation (EC.1) for a problem with horizon T and a policy defined by a parameter $(k+1)$. Namely,

$$\begin{aligned}
\sum_{t=1}^{k+1}R\left(\mathbf{x}_t^{MY}\right)+\sum_{t=k+2}^{T}R\left(\tilde{\mathbf{x}}_{(k+1),t}^{MY}\right) &= R\left(\mathbf{x}_1^{MY}\right)+\sum_{t=2}^{k+1}R\left(\mathbf{x}_t^{MY}\right)+\sum_{t=k+2}^{T}R\left(\tilde{\mathbf{x}}_{(k+1),t}^{MY}\right)\\
&\geq R\left(\mathbf{x}_1^{MY}\right)+\sum_{t=2}^{k}R\left(\mathbf{x}_t^{MY}\right)+\sum_{t=k+1}^{T}R\left(\tilde{\mathbf{x}}_{k,t}^{MY}\right)\\
&= \sum_{t=1}^{k}R\left(\mathbf{x}_t^{MY}\right)+\sum_{t=k+1}^{T}R\left(\tilde{\mathbf{x}}_{k,t}^{MY}\right).
\end{aligned}$$

It remains to show that equation (EC.1) holds for $k=2$. Namely, $\sum_{t=1}^{2}R\left(\mathbf{x}_t^{MY}\right)+\sum_{t=3}^{T}R\left(\tilde{\mathbf{x}}_{2,t}^{MY}\right)\geq R\left(\mathbf{x}_1^{MY}\right)+\sum_{t=2}^{T}R\left(\tilde{\mathbf{x}}_{1,t}^{MY}\right)$. Or equivalently,

$$R\left(\mathbf{x}_2^{MY}\right)+\sum_{t=3}^{T}R\left(\tilde{\mathbf{x}}_{2,t}^{MY}\right)\geq\sum_{t=2}^{T}R\left(\tilde{\mathbf{x}}_{1,t}^{MY}\right). \tag{EC.3}$$

Equation (EC.3) follows from Corollaries 1 and 2. Specifically, inequality (EC.3) compares the revenue collected by two myopic allocations with the same scaled budget update function $\tilde{\mathbf{b}}_{t+1} = \boldsymbol{\phi}\left(\tilde{\mathbf{b}}_t, (1-\gamma)\tilde{\mathbf{q}}\left(\tilde{\mathbf{b}}_t, \tilde{\mathbf{x}}_{k,t}^{MY}\right)\right)$ with initial budgets $\mathbf{b}_2$ and $\tilde{\mathbf{b}}_2$, respectively, such that $\mathbf{b}_2 \geq \tilde{\mathbf{b}}_2$ component-wise. This concludes the proof of equation (EC.1).

Finally, equation (EC.2) follows from equation (EC.1) by comparing the largest left hand side $(k = T)$ with the smallest right hand side $(k = 2)$. This concludes the proof. □

Appendix EC.2: Proofs from Section 5

Proposition 2. *On the instance described in Example 1, with $\gamma > 2$ and $m \geq \max\left\{\frac{1}{\alpha^2}, \frac{2(\gamma-1)}{\gamma-2}\right\}$, both the myopic policy and the L-step look-ahead policy, for any $L \geq 1$, produce revenues that are an $o\left(\frac{\log m}{m}\right)$ fraction of the optimum.*

Proof. We first prove this statement for the myopic policy MY. The proof proceeds in two parts. First, in part 1 we characterize the allocation that the myopic policy induces and upper bound the revenues it can garner over the T time periods. In the second part, we construct a sub-optimal policy and lower bound the revenues it garners. The result then follows by comparing these upper and lower bounds.

Part 1. Assume that at some time t all customers $1 \leq j \leq m$ have their current budget $b_{j,t} = \frac{\gamma}{\gamma+m-1}$. For this budget level, consider an allocation $\mathbf{x}^{\text{uniform}}$ which, for all $j \in \{1, \dots, m\}$, allocates product 1^j to customer j and then fills their remaining budgets with product 2. Finally, $\mathbf{x}^{\text{uniform}}$ fills customer $(m+1)$'s budget with the remaining supply of product 2. Namely,

$$
\begin{aligned}
x_{1^l j,t}^{\text{uniform}} &= \begin{cases} \frac{1}{m^2}, \text{ if } l = j \text{ and } l, j \in \{1, \dots, m\}, \\ 0, \text{ if } l \neq j \text{ and } l, j \in \{1, \dots, m\}, \end{cases} \\
x_{2j}^{\text{uniform}} &= \begin{cases} \frac{\gamma}{\gamma+m-1} - \frac{1}{m^2}, \text{ if } j \in \{1, \dots, m\}, \\ m - \frac{\gamma m}{\gamma+m-1} + \frac{1}{m}, \text{ if } j = m+1. \end{cases}
\end{aligned} \tag{EC.4}
$$

It is easy to verify that $\mathbf{x}^{\text{uniform}}$ is budget feasible for any $\gamma > 2$ and $m \geq 1$.

Applying Lemma 5, where the assumptions that $\gamma > 2$ and $m \geq \left\{\frac{1}{\alpha^2}, \frac{2(\gamma-1)}{\gamma-2}\right\}$ are used, we know that at $t = 1$ with a budget $\mathbf{b}_1$ as defined in Example 1, the myopic policy implements the allocation $\mathbf{x}^{\text{uniform}}$. It is easy to verify that $\phi(b_{j,1}, q(b_{j,1}, \mathbf{x}_j^{\text{uniform}})) = b_{j,1} = \frac{\gamma}{\gamma+m-1}$ for all customers $j \in \{1, \dots, m\}$. Thus, $\mathbf{b}_1$ is a fixed-point under the myopic policy; this implies that MY satisfies $b_{j,t} = \frac{\gamma}{\gamma+m-1}$ for all customers $j \in \{1, \dots, m\}$ and periods $t = 1, \dots, T$. We can then construct an upper bound on J^{MY} by assuming that the myopic policy completely exhausts customer budgets in each period over the horizon T:

$$
J^{\text{MY}}(\mathbf{b}_1) \leq T\left(\frac{\gamma m}{\gamma + m - 1} + 1\right) = 3\log_\alpha(e^{-1})\log m\left(\frac{\gamma m}{\gamma + m - 1} + 1\right) = \Theta(\log m), \tag{EC.5}
$$

where the term $\frac{\gamma m}{\gamma+m-1}$ corresponds to the budgets of customers $1,\ldots,m$, and the 1 corresponds to the budget of customer $(m+1)$.

Part 2. Now, let us consider an alternative policy π^{deplete}, which concentrates all the items on customer 1 for each period t such that $1\leq t\leq T-2$. Namely, for $1\leq t\leq T-2$, π^{deplete} sets:

$$x^{\text{deplete}}_{1^l j,t}=\begin{cases}\frac{1}{m^2}, \text{ if } j=1\\ 0, \text{ otherwise,}\end{cases}$$

and fills the remaining customer budgets with product 2. Clearly, for all customers $2\leq j\leq m$ and periods $t\leq T-2$, $q_{j,t}(b_{j,t},\mathbf{x}^{\text{deplete}}_{j,t})=0$ so that $b_{j,t+1}=\alpha b_{j,t}$. Thus, under the π^{deplete} policy, the budget of each customer $2\leq j\leq m$ at time $T-1$ will be small enough such that the customer can be fully satisfied by only allocating a fraction of item 1^j to her. Namely,

$$b_{j,T-1}=\alpha^{T-2}\frac{\gamma}{\gamma+m-1}=\alpha^{\log_\alpha(m^{-3})-2}\frac{\gamma}{\gamma+m-1}=\frac{1}{\alpha^2m^3}\frac{\gamma}{\gamma+m-1}\leq\frac{1}{\alpha^2m^3}\leq\frac{1}{m^2},$$

where the first inequality follows from $\frac{\gamma}{\gamma+m-1}\leq 1$ for any $m\geq 1$, and the second inequality follows from the assumption $m\geq\frac{1}{\alpha^2}$. Then, at period $T-1$ the policy π^{deplete} sets:

$$x^{\text{deplete}}_{1^l j,T-1}=\begin{cases}0, \text{ if } j=1\\ \alpha^{T-2}\frac{\gamma}{\gamma+m-1}, \text{ if } j\neq 1, l=j\\ 0, \text{ otherwise,}\end{cases}$$

thus obtaining $q_{j,T-1}(b_{j,T-1},\mathbf{x}^{\text{deplete}}_{j,T-1})=1$ and $b_{j,T}=\alpha b_{j,T-1}+(1-\alpha)\geq 1-\alpha$ for all customers $2\leq j\leq m$. Lastly, in period T, π^{deplete} exhausts the budgets of all customers $2\leq j\leq m$ with products of type 2. Therefore,

$$J^{\pi^{\text{deplete}}}(\mathbf{b})\geq(1-\alpha)(m-1)=\Theta(m). \tag{EC.6}$$

Combining equations (EC.5) and (EC.6), we obtain that:

$$\frac{J^{\text{MY}}(\mathbf{b}_1)}{J^*(\mathbf{b}_1)}\leq\frac{J^{\text{MY}}(\mathbf{b}_1)}{J^{\pi^{\text{deplete}}}(\mathbf{b}_1)}=o\left(\frac{\log m}{m}\right).$$

Where the inequality holds since $J^{\pi^{\text{deplete}}}(\mathbf{b}_1)$ is a lower bound on $J^*(\mathbf{b}_1)$.

The proof that

$$\frac{J^{\pi^{L\text{-LA}}}(\mathbf{b}_1)}{J^*(\mathbf{b}_1)}=o\left(\frac{\log m}{m}\right),$$

proceeds similarly and is omitted. The main difference is that we use a different argument to show that starting at $\mathbf{b}_1$, $\pi^{L\text{-LA}}$ allocates uniformly products 1^j to j, thus implying that $\mathbf{b}_1$ remains a steady state budget. This argument is precisely described in Lemma 5. □

Lemma 5. *On the instance described in Example 1, and for $\gamma>2$ and $m\geq\left\{\frac{1}{\alpha^2},\frac{2(\gamma-1)}{\gamma-2}\right\}$, both the myopic and L-step look-ahead policies produce $\mathbf{x}^{uniform}$ as the optimal allocation for $t=1$.*

Proof. We prove, separately for the case of the myopic policy and the L-step look-ahead policy, that $\mathbf{x}^{\text{uniform}}$ is yielded as the optimal allocation on the instance described in Example 1.

Part 1: Myopic policy. Note that at period $t=1$ all customers $1 \le j \le m$ start with equal budget levels. Therefore, by symmetry, the optimal allocation can be either one which allocates each good $\{1^j\}_{1\le j\le m}$ to the corresponding customers $1 \le j \le m$, which is $\mathbf{x}^{\text{uniform}}$, or one which concentrates all these goods to, without loss of generality, customer 1 only. By examination, we can see that at this budget level, $\mathbf{x}_{2j}^{\text{uniform}} = \frac{\gamma}{\gamma+m-1} - \frac{1}{m^2}$ for customers $1 \le j \le m$, and thus $\mathbf{x}_{2(m+1)}^{\text{uniform}} = m - \frac{m\gamma}{\gamma+m-1} - \frac{1}{m} > 0$ for customer $(m+1)$, while all customers $1 \le j \le m$ have their budgets exhausted. Since the price garnered by allocating good 1^j to customer j is higher than the price of allocating some other good 1^l, $l \ne j$, to customer j, then under the concentrated allocation there is less leftover of good 2 to allocate to customer $(m+1)$, and as such $\mathbf{x}^{\text{uniform}}$ garners more revenues in this period than the concentrated allocation. This highlights the purpose of customer $(m+1)$ in Example 1: without this customer, the optimal myopic solution at $t=1$ would be non-unique, with both concentrated and uniform allocations yielding the same objective value, whereas the presence of customer $(m+1)$ solves this issue and guarantees that $\mathbf{x}^{\text{uniform}}$ is the unique myopic allocation.

Part 2: L-step look-ahead. Consider the L-step look-ahead policy with a starting budget $b_{j,1} = \frac{\gamma}{\gamma+m-1}$ for $1 \le j \le m$. Note that the structure of the problem is such that it is always optimal to fully exhaust the budget of a customer $1 \le j \le m$; moreover, since the supply of good 2 is m, it is always possible to exhaust these budgets. Thus, the optimal sequence of allocations produced by $\pi^{L\text{-LA}}$ is the one that maximizes $\sum_{t=1}^{L+1}\sum_{j=1}^{m} b_{j,t}$.

First, note that under any sequence of allocations $\mathbf{z}^1, \ldots, \mathbf{z}^{L+1}$, the budget trajectories in the $L+1$ periods, namely $\mathbf{b}_1, \ldots, \mathbf{b}_{L+1}$, are such that for any j, $b_{j,t} \ge \alpha^L \frac{\gamma}{\gamma+m-1} \ge \frac{\gamma}{2(\gamma+m-1)}$. This implies $\frac{1}{m^2} \le \frac{\gamma}{2(\gamma+m-1)} \le b_{j,t}$, where the first inequality holds for any $m \ge 2$ and $\gamma \ge 2$, and thus $\mathbf{x}^{\text{uniform}}$ is feasible for any $\mathbf{b}_t$ along the budget path. Moreover,

$$U_j^*(b_{j,t}) = \gamma m \frac{1}{m^2} + m \min\left\{\frac{b_{j,t} - \frac{1}{m^2}}{1 - \frac{1}{m}}, \frac{m-1}{m^2}\right\} = \frac{\gamma}{m} + \min\left\{\frac{m^2 b_{j,t} - 1}{m-1}, \frac{m-1}{m}\right\}, \tag{EC.7}$$

and, since $b_{j,t} \ge \frac{\gamma}{2(\gamma+m-1)}$, then $\frac{m^2 b_{j,t}-1}{m-1} \ge \frac{\frac{\gamma m^2}{2(\gamma+m-1)}-1}{m-1} \ge 1 \ge \frac{m-1}{m}$ as long as $\gamma > 2$ and $m \ge \frac{2(\gamma-1)}{\gamma-2}$. Thus, equation (EC.7) becomes

$$U_j^*(b_{j,t}) = \frac{\gamma-1}{m} + 1, \quad \text{for } t \le L+1. \tag{EC.8}$$

Now observe that given any budget level $\mathbf{b}_t$ such that $b_{j,t} \ge \frac{\gamma}{2(\gamma+m-1)}$ and for any feasible allocation $\mathbf{x}$,

$$\sum_j q_{j,t}(b_{j,t}, \mathbf{x}_j) = \sum_j \frac{\gamma m x_{1^j j} + m\sum_{l\ne j} x_{1^l j}}{U_j^*(b_{j,t})} = \frac{\gamma m \sum_j x_{1^j j} + m \sum_j \sum_{l \ne j} x_{1^l j}}{\frac{\gamma-1}{m}+1}.$$

We can cast optimizing $\sum_j q_{j,t}(b_{j,t}, \mathbf{x}_j)$ as the fractional knapsack problem:

$$\max_{\mathbf{x}\geq 0} \frac{\gamma m \sum_j x_{1^j j} + m \sum_j \sum_{l\neq j} x_{1^l j}}{\frac{\gamma-1}{m}+1}$$
$$\text{s.t.} \sum_j x_{1^l j} \leq \frac{1}{m^2}, \quad \forall l.$$

Since we have set $\gamma > 2$, the solution to the knapsack is the one that allocates all the supply of good 1^j to customer j, for each $1 \leq j \leq m$, i.e. sets $x_{1^j j}$ to its maximum feasible value for all j, which is precisely the allocation $\mathbf{x}^{\text{uniform}}$. This implies that, starting at $t = 1$, choosing $\mathbf{x}^{\text{uniform}}$ at $t = 1, \ldots, L+1$ maximizes each $\sum_j q_{j,t}(b_{j,t}, \mathbf{x}_j)$, and consequently $\sum_{t=1}^{L+1} \sum_{j=1}^{m} b_{j,t}$. Thus $\mathbf{x}^{\text{uniform}}$ is the optimal first period allocation for $\pi^{L\text{-LA}}$. $\square$