UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

DYNAMIC PROBLEMS IN REVENUE MANAGEMENT: STRUCTURE AND APPROXIMATION.

TESIS PARA OPTAR AL GRADO DE
DOCTORA EN SISTEMAS DE INGENIERÍA

DANA MARÍA PIZARRO

PROFESOR GUÍA:
JOSÉ CORREA HAEUSSLER

MIEMBROS DE LA COMISIÓN:
SANTIAGO BALSEIRO
OMAR BESBES
MARCELO OLIVARES ACUÑA
GUSTAVO VULCANO

Este trabajo ha sido parcialmente financiado por ANID-PFCHA/Doctorado Nacional/2016 # 21161440

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTORA EN SISTEMAS DE INGENIERÍA
POR: DANA MARÍA PIZARRO
FECHA: 2020
PROF. GUÍA: JOSÉ CORREA HAEUSSLER

DYNAMIC PROBLEMS IN REVENUE MANAGEMENT: STRUCTURE AND APPROXIMATION.

A pesar de que los orígenes del Revenue Management se remontan a su utilización por las compañías aéreas en la era posterior a la desregulación en los Estados Unidos, hoy en día es muy utilizada por diversas industrias como hoteles, empresas de alquiler de automóviles y retailers. Desde entonces, la investigación en esta área creció y se convirtió en una de las ramas más exitosas de la investigación operativa. Cuánto, a qué precio y cuándo vender son las preguntas claves que el Revenue Management pretende responder para maximizar los ingresos esperados de un vendedor. Diferentes supuestos llevan a diferentes modelos y, en consecuencia, surgen nuevos desafíos tanto teóricos como algorítmicos. En esta tesis abordamos algunos de ellos, estudiando desde enfoque teórico de un problema de determinación de precios dinámicos hasta el estudio del gap de optimalidad de algunos problemas de decisión online y la relación entre algunos de ellos.

En el primer capítulo, estudiamos el problema al que se enfrenta un vendedor dotado de una sola unidad de producto para la venta en un horizonte infinito con el fin de maximizar sus ingresos esperados. La empresa se compromete previamente a la función de precio y el comprador es estratégico y tiene una valoración privada por el ítem. Nuestro objetivo es estudiar la importancia, en términos de los ingresos esperados del vendedor, de observar el tiempo de llegada del comprador. En ese sentido, nuestro principal resultado establece que el ingreso esperado cuando el vendedor observa la llegada del comprador es a lo mas aproximadamente 4,91 veces el obtenido cuando el vendedor no observa la llegada del comprador.

En el segundo capítulo, mostramos que si tenemos un mecanismo posted price con una cierta garantía de aproximación, podemos obtener prophet inequality (en el mismo escenario) con la misma garantía de aproximación. Este resultado, junto con el trabajo de Chawla et al. [41], implican que el problema de diseñar mecanismos de posted price es equivalente al de encontrar reglas de parada para el problema de tiempo de parada óptimo.

Por último, en el tercer capítulo presentamos una clase general de problemas de decisión online, llamado Dynamic Resource Constrained Reward Collection (DRCRC), que comprende varios problemas estudiados por separado en la literatura. Debido a la dificultad de resolver el problema de forma óptima, es habitual desarrollar heurísticas fáciles de resolver y que sean una buena aproximación para el ingreso óptimo esperado. Así, estudiamos la pérdida de ingresos bajo una heurística estudiada en la literatura, proporcionando una única prueba de la pérdida de ingresos para los problemas incluidos en la clase DRCRC, recuperando resultados existentes para algunos problemas y obteniendo nuevos para otros.

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTORA EN SISTEMAS DE INGENIERÍA
POR: DANA MARÍA PIZARRO
FECHA: 2020
PROF. GUÍA: JOSÉ CORREA HAEUSSLER

DYNAMIC PROBLEMS IN REVENUE MANAGEMENT: STRUCTURE AND APPROXIMATION.

Although the origins of Revenue Management date back to its use by airlines in the post-deregulation era in the U.S., nowadays is heavily used by a variety of industries such as hotels, car rental companies and retailers. Since then, research in this area grew and became one of the most successful branches of operation research. How much, at what price and when to sell are the key questions that Revenue Management aims to answer in order to maximize the expected revenue of a seller. Different assumptions carry to different models and consequently, new theoretical and algorithmic challenges arise. In this thesis, we address some of them, going from a theoretical approach of a dynamic pricing problem to the study of the optimality gap for some online decision problems and the relation between some of them.

In the first chapter, we study the problem faced by a seller endowed with a single unit for sale over an infinite time horizon in order to maximize her expected revenue. The firm pre-commits to the price function and the buyer is strategic and has a private value for the item. Our goal is to study the importance, in terms of seller's expected revenue, of the observability of the buyer's time arrival. Our main result states that, in a very general setting, the expected revenue when the seller observes the buyer's arrival is at most roughly 4.91 times the expected revenue when the seller does not know the time when the buyer arrives.

In the second chapter, we show that if we have a posted price mechanism with a certain approximation guarantee, this can be turned into a prophet inequality (in the same setting) with the same approximation guarantee. This result, together with the work by Chawla et al. [41], imply that the problem of designing posted price mechanisms is actually equivalent to that of finding stopping rules for a prophet.

Finally, in the third chapter we introduce a general class of online decision problems, namely Dynamic Resource Constrained Reward Collection (DRCRC) that comprises several problems studied separately in the literature. In particular, the class of DRCRC admits as special cases the classes of network revenue management problems, dynamic pricing problems, online matching problems, to name a few. Due to the difficulty of solving the decision-maker problem optimally, it is usual to develop heuristics to approximate the optimal expected revenue. Thus, we study the revenue loss under a well studied certainty equivalent heuristic, providing a unifying proof for the revenue loss for DRCRC, recovering some existing results for some problems and obtaining new ones for others.

"I was taught that the way of progress is neither swift nor easy."
Marie Curie.

Aknowledgements

First of all, thanks to José, who was an excellent advisor since the beginning of my Ph.D. I would like to thank you not only for the way you share your knowledge and passion for research, but also for your commitment to your students and projects. You're definitely the best advisor I could have had! Thank you for trusting on me!

Secondly, I would like to thank Gustavo Vulcano, with whom I have worked most of my years of research in the Ph.D, and from whom I have learned a lot, not only from his academic knowledge but also from the way he works. Thank you for your advices, your work, and especially for your patience! Thank you also for opening the doors of the Universidad Torcuato Di Tella, you really made me feel at home!

Thirdly, I would like to thank Omar Besbes and Santiago Balseiro, who have given me the opportunity to work with them and spend some time at Columbia University. They are definitely the perfect match: Omar always has the right intuition for everything and a list of papers related to each thing in his head, Santiago is always in every detail and always has the answer to the issues, knowing not only which theorem is useful, but in which book and on which page to find it. I definitely learn a lot from both of you! Thanks a lot for that!

I would also like to thank my co-authors Victor Verdugo and Pato. It was a pleasure to work with you! Thanks Victor because I always made you crazy with questions and advices and you were (and you are) always there, like a good academic brother.

On the other hand, I want to thank to Victor Bucarey, who is somehow "guilty" of my passage through this Ph.D. Thanks a lot for the long talks, for always giving me support when I needed it, for the company during most of these years in Chile and for being a great friend!

Thanks to all my colleagues at DSI, because without the lunches, the seminars and the outing plans, the experience would not have been the same. In particular, thanks to those who have become more than colleagues: Verito, Richi, Seba, Coni, Renny, Edu1, Edu2, Vale, Colombia, Feña. Your friendship and the infinite memories with you is one of the best things I will take with me from Chile.

Thanks also to the república team: “minions de José”.. definitely the most nerdy talks during my Ph.D. were during lunches with you!

I also want to thank the people who supported me and were part of my path from the other side of the mountain range. First of all, thanks to the good crazy girls: Peque, Viky and Sole, because without them my life would not be the same. Thank you because distance showed us that true friendships not only survive to it, but get stronger with every difficulty.

Thank you to my brothers and sisters: Aixa, Matu, Maqui, Yair and Alejo, who were and are always there, in the best and in the worst moments; and to my parents, who have always supported and accompanied me in each and every decision I have taken.. you are definitely my inspiration.

Thanks to Franco and Sasha who, during each visit to Argentina, each kiss, hug and “I love you” recharged me with new energy to go on. Thank you for saying “chau Dana!!!” to every plane that passes by, it keeps me close despite the distance. I love you both to infinity and beyond.

Thanks to all my family and friends, without their unconditional love I would not have made it.

Contents

Introduction

Revenue Management can be defined as the art of being efficient in the allocation of goods and services with limited inventory over a given selling horizon. Due to consumers in a market are heterogeneous, and not all have the same willingness to pay for the product or service the company offers, it may be optimal for the seller to offer the product to different segments at different prices. For instance, one way to do this is through discounts, offering dynamic prices (i.e., prices changing over time), or simply offering higher quality service at a higher price. This requires decisions to be made about how much product to sell to each type of costumer and at what price to do so, which is the basis of revenue management.

Its origins date back to the seventies with a work by Littlewood, who was the first to propose a solution method for the seat inventory control problem for a single leg flight with two fare classes. Basically, this problem from the airline industry states as follows: there are a finite number of seats of a single leg flight to be allocate to customers who arrive over time. Two fare classes are considered, with levels f_l and f_h, with $f_l < f_h$. Once a booking request is received, the airline should decide whether to accept or reject it. The goal is to search for a booking control policy maximizing the expected revenue for the whole selling horizon. Note that in general it is not easy to find it because if the decision-maker rejects a lot of demand of low fare class, there could be empty seats at the end of the selling horizon and therefore the policy is sub-optimal; on the contrary, if he accepts too many requests for low fare class at the beginning of the selling horizon, it may be that later high fare class demand must be rejected due to lack of capacity, losing revenue. To solve this problem, Littlewood assumed that demand for low fare class comes before demand for high fare class and his policy consists on stop accepting the low fare booking requests once revenue from selling another low fare seat is exceeded by the expected revenue of selling the same seat at the higher fare. That is, demand for low fare class should be accepted as long as $f_l \geq f_h \mathbb{P}(D_h > C)$, where D_h represents the demand for high fare class and C is the inventory left. This leads to the well known *Littlewood's Rule* that characterizes the optimal amount of inventory to be reserved to satisfy demand from high fare class.

Few years later, as response to the deregulation of U.S. domestic and international airlines—which led to the arrival of competition on the market— American Airlines implemented the Littlewood Rule and pioneered the real application of revenue management in the world. Since then, the airline revenue management problem has received a lot of attention throughout the years and continues to be of interest. Furthermore, nowadays revenue management is one of the most successful areas of operations management, both for its usage in practice by several industries and for its theory developments.

One of the cornerstones of the revenue management on which a large number of scientists started working during the last few decades is *dynamic pricing*. Dynamic pricing aims at maximizing profits by dynamically changing offer prices within the selling horizon in order to optimally exploit changes in demand or competition-related conditions. Intertemporal price discrimination, which is regarded as the basis of dynamic pricing, has been studied since the seminal work of Coase [42]. He considers a monopolist who sells a durable good to a large set of customers with different valuations and analyzes how the seller has to price the product in a way that at the beginning the price is high enough to capture high-valuation customers and then sequentially reduce the price to capture customers with smaller valuation. If this strategy works, it results in extracting a large fraction of the consumer surplus. However, Coase argues that if high-valuation customers anticipate that prices will decrease, they would wait for a lower fare. This, in equilibrium, will lead the seller to offer the product at marginal cost. This result does not hold when supply is infinite or the good is perishable because consumers may not have the incentive to wait for the lower price.

How to optimally adjust prices periodically is a question that can be studied under several model variants, including number of items to sell (single vs. multiple), the relative position of the seller in the market (monopolistic vs. oligopolistic), the degree of rationality of the consumers, the seller's ability to change prices over time, and the length of the horizon (finite vs. infinite). Despite the progress in studying different settings of the problem, most existing literature relies on the seller's ability to know the buyer's arrival to the market. When the seller can observe the arrival of the buyer, she can make the price function contingent on the buyer's arrival time, improving her profits. However, it may not be realistic in some contexts, such as online marketplaces, and this motivated us to think about what is the additional rent that the seller can obtain by having the ability to observe the arrival of the customers. To this end, in the **first chapter** we study how important is to observe the arrival time of the buyer in terms of the seller's expected revenue in a simple model. More specifically, if we defined the *value of observability* as the worst case ratio between the expected revenue of the seller when she observes the buyer's arrival and that when she does not, our main states that, in a very general setting, the value of observability is at most 4.911. To show that, we fully characterize the observable setting and use this solution to construct a random and periodic price function for the unobservable case.

Broadly speaking, the questions that are addressed in revenue management are also studied in the *Mechanism Design* literature, which aims to find the rules of a system so that, in equilibrium, the desired objective of the decision-maker is optimized. One of the most relevant mechanisms developed in the last decade is the Posted Price Mechanism (PPM). In this setting, consumers are faced with take-it-or-leave-it offers, and therefore strategic behaviour of consumers simply vanishes. Each customer buys the offered item if and only if his valuation is not lower than the observed price. Because the problem of computing the price seems to be much simpler than optimal auctions, online sales have been moving from an auction format to a posted price format, which has received significant attention in the last decade from researchers on computer science trying to understand the approximation guarantees that can be obtained through this type of mechanism.

In that way, Hajiaghayi et al. [79] and later Chawla et al. [41] established a surprising connection between sequential posted price mechanisms and prophet inequalities, an old theory arising in optimal stopping. Implicitly they show that any prophet type inequality can be turned into a posted price mechanism with the same approximation guarantee. As a consequence, most follow up work in the field concentrated on prophet inequalities and then applied the obtained results to sequential posted price mechanisms. A question that remained unsolved is whether the converse also holds, that is, if we have a sequential posted price mechanism with a certain approximation guarantee, can this be turned into a prophet inequality (in the same setting) with the same approximation guarantee?

In the **second chapter** of the thesis we answer this question on the positive implying that the problem of designing sequential posted price mechanisms is actually equivalent to that of finding stopping rules for a prophet. Our reduction is robust to multiple settings including having matroid constraints and downward-closed feasibility constraints, or different arrival orderings such as deterministic, random, or worst case. The crux of our analysis is a new Lemma in mechanism design— that we believe may find applications that go beyond the scope of this thesis— stating that for any random variable X, there exists another random variable Y whose ironed virtual valuation is distributed as X. As a consequence of our main results we obtain improved lower bounds for the performance of sequential posted price mechanisms in the bayesian single-parameter setting by carrying the lower bounds on i.i.d. prophet inequalities of Hill and Kertz [81] to the pricing setting.

PPM is one of the problems studied in the literature of *online optimization*, which is widely used to afford problems where the decision-maker has to choose a feasible action immediately after an arrival and there is uncertainty— or there is no information— about the future. Due to the complexity of computing the optimal policy, even for problems with moderate size, researchers work to find heuristics that are easy to implement and have a good revenue performance compared to the offline optimum, that is, with the optimum obtained if the problem is solved with complete information.

These class of problems have received special interest in the last decades both in the Revenue Management and in Computer Science community. One drawback of the most existing literature is that they study each of the problems separately. However, it is possible to define a general model that allow us to think each of them as particular cases of it, and this is part of the contributions of the **third chapter** of the thesis. Explicitly, we define a class of problems: dynamic resource constrained reward collection (DRCRC), which comprises a lot of the online decision problems studied in the literature, and thus the model involved is more general. In words, a problem in DRCRC states as follows: opportunities arrive over a finite time horizon. Upon arrival, a decision-maker must choose an action by relying only on present and past information, whereas future information is uncertain. Such actions have associated some resources consumption— with finite initial inventory— and a reward collection. The goal of the decision-maker is to select actions maximizing his total expected reward subject to the resource consumption constraints. We consider a certainty equivalent heuristic and we study its performance for the DRCRC class of problems, yielding a unifying proof to the perfomance of the heuristic for the studied problems that are special cases of the class DRCRC.

Most of the material in this thesis has been or will be published. The first chapter is based on joint work with José Correa and Gustavo Vulcano [47], whose results have recently appeared in the proceedings of the 21*st* ACM Conference on Economics and Computation. The material in Chapter 2 is based on joint work with José Correa, Patricio Foncea and Victor Verdugo [45], results published in Operations Research Letters. Finally, the material in Chapter 3 is based on a working paper with Omar Besbes and Santiago Balseiro [19].

Each chapter will be organized as follows: first we present an introduction to the problem and the main contributions of the chapter. Then, we carry out a brief review of the most relevant literature regarding our goals and after that we make a more precise definition of the model/problem to be studied during the chapter. Finally, we present the results. In order to facilitate the reading and understanding of the thesis, we refer the proofs in the appendix of each chapter.

Chapter 1

On the observability of the arrival time of consumers in a dynamic pricing problem[1]

1.1 Introduction

In recent years we have witnessed an enormous amount of work in dynamic pricing and dynamic mechanism design. Driven by the increasingly important online marketplaces, the area has been particularly active in Economics, Operations Management and Computer Science. Although the borders are blurred, often research in operations management deals with finding optimal or approximately optimal dynamic pricing mechanisms (see, e.g., [26, 38, 69]), in economics the central interest is to find optimal dynamic mechanisms (see, e.g., [29, 114]) which may involve departing from basic pricing schemes, while in computer science the interest is in designing simple mechanisms which are approximately optimal (see, e.g., [28, 41, 45]).

One drawback of part of the literature is the underlying assumption that the seller is informed of the buyers' arrivals. This assumption allows the seller to update the pricing/mechanism when observing a new arrival. In some contexts, such as in online marketplaces, it may be difficult for the seller to distinguish interested buyers from other traffic on the website and therefore assuming that the seller observes the buyers' arrivals may not be realistic. The extent to which this observational ability produces additional rents to the seller is the main subject of this chapter.

Specifically, we consider a simple, yet fundamental, model in which one seller interacts with a single buyer. The seller holds a single item whose value is normalized to zero, while

[1] This chapter of the thesis is based on joint paper with José Correa and Gustavo Vulcano [47]

the buyer has a private random valuation for it. The buyer arrives according to an arbitrary distribution over the nonnegative reals. As usual in the literature, both the buyer and the seller discount the future but they do it at different rates, the buyer being more impatient than the seller. The goal of the seller is to set up a price function so as to maximize her expected discounted revenue. On the buyer's side, upon arrival, he observes the price function and decides to buy at the time that is most profitable for him. The ability of the seller to observe the buyer arrival (or not) determines two different situations. In the *observable* case, the seller observes the buyer arrival and thus the price function she sets may be dependent on it. In the *unobservable* case this ability is absent and therefore the seller has to set a price curve from the beginning of the selling horizon only knowing the arrival distribution. These two scenarios naturally lead to define the *value of observability* for a given instance of the problem as the ratio between the revenue of the seller in the observable case and that in the unobservable case. A particular instance of the problem is defined by the arrival distribution of the buyers and their valuation distribution, and the discount rates for both the buyer and the seller. Then, the more general *value of observability* (VO) is defined as the supremum of the corresponding instance-specific VO taken over all possible distributions and discount rates, which corresponds to the worst-case ratio between the revenue of the seller in the observable and unobservable cases. The focus of our work is to bound this worst-case ratio.

There are two equivalent interpretations for this model that are worth highlighting. First, on the demand side, the model with a single buyer could be equivalently interpreted as having a continuum of buyers with total mass equal to 1. In the observable case, the infinitesimal mass of buyers is represented by the probability density function (pdf) of the buyer's private valuation. This interpretation is extended by also accounting for the pdf of the buyers' arrival distribution in the unobservable case. On the supply side, we assume a unit supply which is infinitesimally partitioned so that it can be taken as unlimited.

Second, our VO result can also be interpreted as the *price of discrimination*. To see this, consider the infinitesimal buyer view of our model just introduced, where the seller sets a *personalized price curve* for each arriving buyer so that the total expected discounted revenue she obtains is the same as that achieved in the observable case. On the other hand, if the seller does not have this power, she should offer the same pricing policy for all customers since the beginning of the selling horizon. The latter problem is exactly the same as the unobservable case described above. Therefore, if we define the *price of discrimination* as the additional rent the seller can obtain by tracking each buyer arrival time and posting a personalized price curve, it becomes equivalent to the value of observability. Thus, we are also providing a bound for the price of discrimination, i.e., for the additional rent the seller can obtain by offering a personalized price curve for each customer.

The value of price discrimination is being studied by researchers in operations due to its practical interest. Although most of them are focused on how to do price discrimination, a recent work by Elmachtoub et al. [56] studies when doing price discrimination is worthwhile and when it is not. Specifically, they provide lower and upper bounds on the ratio between the revenue achieved from charging each costumer his own valuation and the revenue obtained through a single price strategy and they also compare the profit obtained when the seller observes some information before fixing the pricing policy (but not the buyer valuation) with the one earned by each of the two strategies described above.

Motivating Example

A key difficulty in evaluating the value of observability is that the unobservable case is typically very hard to solve and standard approaches to tackle dynamic pricing or mechanism design problems based on optimal control fail.

To better grasp this difficulty and the difference between the observable and unobservable cases let us describe a quick example. Take a buyer with valuation uniformly distributed in $[0, 1]$ and arrival time distributed as an exponential with mean 1. Also assume the seller discount rate is 1 while that of the buyer is extremely large (so that in the end the buyer is myopic, he will buy as long as the price is below his valuation). Then, if the seller can observe the buyer's arrival in our dynamic pricing setting, she will start pricing at 1 and then decrease the price suddenly in a continuous fashion until hitting the customer valuation, where the transaction is executed. In this way, she will be extracting all the consumer surplus, with expected value 1/2. Thus, in expectation, the seller gets $\int_0^\infty (\mathrm{e}^{-t}/2)\mathrm{e}^{-t}\mathrm{d}t = 1/4$ (here, the first e^{-t} represents the discounting and the second e^{-t} represents the density of the exponential).

On the other hand, in the unobservable case, if we assume that the seller needs to set a decreasing price function then the problem is relatively easy to solve. Indeed, the seller would need to maximize, over all decreasing functions $p(\cdot)$, the quantity $\int_0^\infty (\mathrm{e}^{-t}(1 - p(t)) - (1 - \mathrm{e}^{-t})p'(t))\mathrm{e}^{-t}p(t)\mathrm{d}t$. Note that for a decreasing $p(t)$, trade occurs between t and $t + \mathrm{d}t$ if either the buyer arrived in that interval and his valuation is above $p(t)$ (hence the term $\mathrm{e}^{-t}(1 - p(t))$), or the buyer arrived before t and his valuation is between $p(t)$ and $p(t + \mathrm{d}t)$ (hence the term $-(1-\mathrm{e}^{-t})p'(t)$). In both cases the discounted revenue for the seller is $\mathrm{e}^{-t}p(t)$. The solution of this problem turns out to be $p(t) = \mathrm{e}^{-t}$, which results in an expected revenue of 1/6. Overall the ratio of the revenues between the observable and non observable cases is 3/2, and therefore this suggests that $VO \geq 3/2$. However, the seller's strategy space is richer than that of decreasing price functions. Suppose that she splits the time horizon into short intervals of length ε and considers a periodic price function that sets price 1 for the first $\varepsilon - \varepsilon^2$ time units of each interval and a quickly decreasing price (from 1 to 0) in the last ε^2 time units of each interval. As the buyer is myopic he will buy at the first point in time in which the price is below his valuation and since ε is very small the probability that the buyer arrives when the price is 1 is close to 1. Thus, even in the unobservable case the seller is able to obtain a revenue arbitrarily close to 1/4.

Furthermore, when the discount rate of the buyer is not too large strategic behavior comes into play, which adds an additional layer of difficulty in formulating the problem, as we discuss in Section 1.4.2. It should be noted that if both the buyer and seller discount at equal rates then the optimal pricing function is simply constant in both the observable and unobservable cases, therefore strategic behavior vanishes, there is no delay on trade, and the resulting expected revenues for the seller are equal. Thus we assume throughout that the buyer's discount rate is strictly larger than that of the seller.

1.2 Contributions

We summarize contributions of Chapter 1 below.

Firstly, we revisit the observable case. Although this problem is far from new, we write the seller's problem as an optimal control problem. Due to this problem is difficult to solve, we propose a relaxation by computing the first order condition of the equilibrium constraints. In fact, we show that the problems are equivalent and we use Euler-Lagrange optimality conditions to derive a characterization of the optimal pricing policy as the solution of an ordinary differential equation. We also prove a key result (Lemma 1) establishing that in the optimal pricing the seller extracts a constant fraction of the total revenue within a short time, that solely depends on the seller's discount rate.

Secondly, we turn to study the unobservable case. Unfortunately, this problem is much harder to analyze and obtaining an explicit solution seems hopeless. However, we described and model the problem under some assumptions over the primitives of the problem. It is worth mentioning that for the main purpose of the work it is enough to to exhibit a pricing policy that can recover a constant fraction of the revenue of the optimal solution in the observable case.

Finally, we prove that for arbitrary arrival and valuation distributions of the buyer and arbitrary discount rates of both the seller and the buyer, the *value of observability* is bounded above by a small constant. To this end, we search for a particular pricing policy for the unobservable case that allows us to compare the seller's expected revenue under this policy with the optimal expected revenue for the observable case. The main idea is to use the solution of the observable case and try to repeat it over time to contract a periodic price function. Of course this is not possible since already that solution takes infinite amount of time to implement. Thus the aforementioned key result comes into play and allows us to do this repeated pricing within small time windows. The second obstacle is that we should be careful with the buyer's strategic behavior. To avoid this issue, we simply introduce *empty space*, say by using a very high price, before each application of the optimal observable pricing so as to make a buyer, arriving within this empty space, behave as in the observable case. Again, this comes at a loss of a constant fraction of the revenue. Finally, a difficulty arises since the arrival distribution might now impose a lot of weight in regions where our price is too low. To overcome this we apply a random shift to our price curve which allows us to treat the buyer's arrival time as if it were uniform on a given interval. Ultimately, by carefully dealing with these three obstacles we are able to show that the proposed pricing scheme obtains an expected revenue of at least a fraction $1/4.911 = 0.203$ of the optimal revenue in the observable case.

For the special and relevant case of valuation distributions having monotone hazard rate, which includes several of the standard distributions, we show that the situation is much simpler. Indeed it is enough to consider a fixed price curve (i.e., the price is constant over the whole period) to recover a fraction 1/e of the revenue in the observable case. We further note that fixed pricing cannot guarantee a constant in general.

This result is somewhat surprising because of several factors: (i) the generality of the model; (ii) the bound is totally independent of the model primitives; and (iii) simple pricing strategies, such as fixed pricing, fail to guarantee a constant bound. We also note that our result is robust to the distribution of arrivals. Indeed, even if the arrival time of the buyer was chosen by an adversary that knows the price function of the seller (but does not know the realization of the random shift) then our bound on the VO still applies.

Let us highlight that we also provide a lower bound for the VO by solving, numerically, the unobservable case for a particular tractable instance.

1.3 Related literature

Although the literature on dynamic pricing is extensive, next we present a brief review of the literature, with more emphasis on those related to the problem considered in this chapter. For further reading see [13, 15, 27, 65, 75, 124, 129, 136].

Typically, the literature studies a game between a seller and one or more buyers. The seller owns an item or a set of items, and the buyers have private valuations for them. The game takes place over a time interval since either the buyers arrive over time or since the buyers and the seller discount the future differently, which implies that there is delay on trade. Then, there are several features that characterize a pricing problem. For instance, whether there is a single of multiple buyers and sellers, the inventory to sell, the information about the valuation for the items, the length of the selling horizon, how buyers arrive to the market, how do the decide when to buy and how do both buyers and sellers discount the future, are some of them. In what follows, we make a small survey of the related literature for some different model variants.

Almost all papers consider a monopoly. By avoiding competition between multiple sellers, the authors significantly reduce the mathematical complexity of their models, which already contain the interaction between customers and sellers and between customers themselves. Some papers in which a monopoly is considered are [14, 48, 46, 92]. Like those, in this chapter we also consider that there is only one seller. However, an oligopoly is generally a more realistic setting. Liu and van Ryzin [100] consider an oligopoly and point out that competition reduces the market power and therefore the profits of each individual seller.

Although Aviv and Pazgal [14] and Elmaghraby et al. [57] consider a monopoly, they examine oligopolies with respect to their proposals for potential model extensions.

In relation to the type of consumers, it is usual to consider myopic (see, e.g., [6, 86, 101]), or strategic (see, e.g., [14, 31, 127]) consumers. Myopic consumers are those who purchase at the first time they are able to achieve a positive surplus. Considering this kind of consumers, future prices have no influence on purchasing decisions. On the other hand, strategic consumers decide when to buy in order to maximize their surplus, even though that may imply waiting to buy. That is, they take into account all future purchase opportunities and delay making a purchase if necessary in order to achieve a higher surplus. It is worth mentioning that some authors consider the presence of both myopic and strategic consumers in their model (see, e.g., [92, 135]). Motivated by the presence of strategic consumers in the market and the related recent literature, our model follows this trend, with a single strategic consumer.

Regarding the arrival of consumers, some papers assume that it is simultaneous at the beginning of the selling horizon. This is mainly because allows to model some applications, as well as it makes the problem mathematically more tractable (see, e.g., [25, 64, 99, 48, 92]). For instance, Gallego et al. [64] presents a model in which a monopolist wants to sell a fixed inventory of a product to consumers arriving at the beginning of the two selling periods (the market size is fixed) with a decreasing willingness to pay for the product. The seller commits to a markdown price path for the two periods and, given that, each consumer determines in which (if either) of the two period will purchase. They prove that, in equilibrium, a single-price policy is optimal if all consumers are strategic and the seller knows the demand. Without any of these conditions, it can be optimal to implement a two-price markdown policy. On the contrary, other authors model a dynamic pricing problem with customers arriving sequentially during the selling horizon. In this case, customers typically arrive following a Poisson process (see, e.g., [14, 57, 31]). However, there are papers considering a different sequential arrival of consumers. For example, Su [127] considers customers being infinitesimally small and that they arrive continuously according to a deterministic flow of constant rate. Wu et al. [133] consider a market size as a random variable with a general distribution, while Gershkov et al. [69] considers that the arrivals are described by a Markov counting process. In particular, we consider that the buyer arrives according to a known distribution.

About how to discount the future, some of the literature only considers discount factor on the customer side. For instance, Correa et al. [31] and Kremer et al. [92] assume that all strategic consumers have an identical discounting factor. That is, the buyer's surplus is discounted by an identical rate for all buyers. The higher a discount factor is, the more impatient the buyers are. On the other hand, Aviv and Pazgal [14] and Cachon and Swinney [35] consider a model where only discount the valuation over time, thus resulting in minor changes in the model and results. On the other hand, it is possible to include discounting on the part of sellers and customers in their models. Instead of assuming that both seller and buyers have the same rate discount (see, e.g., [25, 68, 69]), we consider that buyer's rate discount is grater than the seller's one, which means that the buyer is more impatient than the seller. Regarding that, Osadchiy and Vulcano [112] find that the most beneficial scenarios occur when the seller's discount factor is lower than the customer's and therefore, the seller can take advantage of the customer's impatience. Others papers in which both seller and buyers have no the same rate discount are Mantin and Granot [105] and Correa et al. [31]. It is worth mentioning that some authors model the impatience of buyers by considering a deadline to make a decision of whether to buy or not and when, or an explicit waiting cost (see, e.g., [107, 113, 127]).

Another key feature of our work, is that we study a pricing problem in a continuous time setting. The literature considering continuous time pricing models is rather limited (see [31, 67]), and most of the work is over discrete time, e.g. [26, 46, 48].

Despite the fact that there are a large number of recent scientific papers in dynamic pricing, we are interested in studying the importance, in terms of seller's expected revenue, of the observability of the buyer's time arrival to the market, which is a question that it has not been studied yet. In particular, we bound the additional rent the seller can obtained when she is able to observe the arrival time of the buyer before offering the price curve. To this end, we define what we called *observable case*—the seller observes the buyer's arrival and then set a price curve maximizing her expected revenue—, and the *unobservable case*—the seller is not able to observe the arrival and therefore she should fix the prices since the beginning.

Regarding the observable case, this problem is far from new and indeed already Stokey [126] notes that intertemporal price discrimination happens only due to the difference in discount rates. Later, Landsberger and Meilijson [96] precisely show that this price discrimination through time is optimal, while Shneyerov [125] considers the situation in which there are multiple units to sell. These both works studied the problem from a mechanism design approach. The observable case has been studied also from a pricing approach by Wang [132]. As in the latter, we take a pricing approach (rather than a mechanism design) which, as usual in this literature, allows us to write the seller's problem as an optimal control problem and furthermore to fully characterize its solution. Although Shneyerov [125] considers a very similar situation to ours, they characterize the optimal price function through the maximum principle and, therefore, it involves a rather complicated hamiltonian. On the contrary, our approach is simpler, and based on the Euler-Lagrange optimality conditions we can derive a simpler ODE which we prove has a unique solution.

Although the unobservable case has not been studied in the literature in the context we consider, some papers does not assume complete information about the arrival time. For instance, Caldentey et al. [38] consider a setting where the buyers are characterized not only by their valuation but also by their arrival time. However, their approach differ from ours because their goal is to characterize optimal price curves minimizing the worst-case regret of the seller. The setting they consider is also different: both the buyer and the seller—in fact they first consider only one buyer but then do the analysis for more buyers—discount at the same rate. Furthermore, they consider a finite horizon and both myopic and strategic consumers—in the case of one buyer, they consider that he may be either myopic or strategic. Another related paper is the one by Bergemann and Strack [21], who considers a setting where the arrival time and the valuation is private information of each buyer and unobservable to the seller. The main difference with our model is that they consider both the seller and the buyers discount the future at the same rate, which is necessary to their approach to work.

Organization

We start with the precise model description in Section 1.4, including describing the buyer's problem and the seller's problem in Section 1.4.2. This latter section includes the formulations of the standard observable case and the more challenging unobservable case. Both cases are later analyzed in detail in Sections 1.5 and 1.6, respectively. Finally, the bounds for the VO are established in Section 1.7.

1.4 Model description

We study the problem faced by a firm (seller) endowed with a single unit for sale over an infinite time horizon. The value of the item for the seller is normalized to zero. We take a revenue management (RM) point of view and assume that the seller cannot replenish this unit throughout the selling horizon. On the demand side, a single consumer will arrive at a time that follows a cumulative distribution function (cdf) $G : [0, \infty] \to [0, 1]$ and density g. The buyer has a private valuation v for the item with cdf $F : [0, \bar{v}] \to [0, 1]$ and density f. Both G and F are common knowledge. As it is standard in the literature we can equivalently think that the seller has unlimited supply, and that on the consumer side we have a mass of consumers with arrivals distributed as G and valuation distributed as F.

The interaction between the seller and the buyer is formalized as a Stackelberg game in which the seller is the leader and pre-commits to a price function $p(t)$ over time in order to maximize her expected revenue. The buyer is the follower and has to decide whether and when to purchase the item, given the price function set up by the seller.

We discuss two possible variants of this problem. In the *observable case*, the seller is able to track the buyer's arrival time τ and from that moment onwards she commits to a price function $p : [\tau, \infty] \to [0, \bar{v}]$. In the *unobservable case*, the seller does not see the buyer's arrival time (although she does know the arrival time distribution G) and since time 0 she commits to a price function $p : [0, \infty] \to [0, \bar{v}]$.

Even though for the ease of exposition the game between the seller and the buyer is presented as if the seller were to announce the price function in a first stage, and the buyer were going to decide *if* and *when* to purchase in the second stage, strictly speaking, the game can also be described as a simultaneous game with no need of precommitment since the calculation of the price function and the timing of the buyer's purchase decision are based on common knowledge information.

For technical reasons, in both cases we impose the mild condition that the price function p is lower semi-continuous and differentiable almost everywhere[2]. In what follows, we introduce the buyer's and the seller's problems, as well as some preliminary definitions and results.

1.4.1 The buyer's problem.

When the buyer arrives, he observes the price function for all future times and decides whether and when to buy in order to maximize his utility. We assume that the consumer is forward-looking and sensitive to delay, and denote by $U(t, v)$ the quasilinear discounted utility function of a consumer with valuation v purchasing at time t. In particular, we consider an exponentially discounted utility function: $U(t, v) = \mathrm{e}^{-\mu t}(v - p(t))$, where $\mu > 0$ is the discount factor.[3]

Note that as $t \to \infty$, $U(t, v) \to 0$, so that the buyer eventually purchases the item as long as $v > p(t)$, for some t. Given a price function $p(t)$, a forward-looking buyer arriving at time τ with valuation v solves:

$$[BP] \qquad \max_{t \geq \tau} U(t, v).$$

It may be possible that this problem has multiple solutions, and to avoid ambiguity we will further assume for convenience that the buyer purchases the item at the earliest time maximizing his utility. We define the auxiliary function $\phi : [0, \infty) \to [0, \overline{v}]$ as:

$$\phi(t) = \inf\{v : U(t, v) \geq U(t', v),\ \forall t' \geq t\},$$

that represents the minimum valuation that the buyer must have in order to buy at time t and not later, and it is defined irrespective of the buyer's arrival time $\tau \leq t$. In other words, the function ϕ defines a threshold in the sense that if a buyer with valuation v buys at time t, then a buyer with valuation $v' > v$ buys at the same time and not later[4]. Based on ϕ, we are able to describe the equilibrium conditions for the buyer purchasing behavior and used them to formulate the seller's problem.

[2]Due to this assumption the seller could potentially lose at most a negligible extra revenue and therefore it does not affect our results. Moreover, the lower semi-continuity is necessary to ensure that the buyer's problem can always be solved.

[3]This intertemporal utility function discounts the buyer's payoff from time zero, and it is without loss of generality for the sake of characterizing an optimal policy. That is, if the buyer purchasing at time t only incurs the disutility for waiting from his arrival time τ, then the utility function $U(t, v)$ would only be affected by a fixed constant: $U(\tau, t, v) = \mathrm{e}^{-\mu(t-\tau)}(v - p(t)) = \mathrm{e}^{-\mu\tau}U(t, v)$.

[4]To see this, knowing that $v = \phi(t)$, we have $U(t, v) \geq U(t', v)\ \forall t' \geq t$. Now, consider a buyer with valuation $v' = v + \varepsilon, \varepsilon > 0$. By simple algebra we have $U(t, v') = U(t, v) + \varepsilon \mathrm{e}^{-\mu t} > U(t', v) + \varepsilon \mathrm{e}^{-\mu t'} = U(t', v')$, i.e., $U(t, v') \geq U(t', v'), \forall t' \geq t$. Thus, the purchasing time of buyer v' cannot be later than t.

1.4.2 The seller's problem

The seller's problem is to select a price function to maximize her expected revenue, taking into account the forward-looking behavior of the buyer.

Observable arrival case

In this situation the seller observes the arrival time τ of the buyer and therefore sets a price function p defined over $[\tau, \infty)$. For now, we will pretend that the buyer arrives at time zero, i.e., we initially assume that $\tau = 0$.

Given the threshold function ϕ induced by the price function p, a buyer with valuation v will purchase at the first time $t \geq 0$ satisfying $v \geq \phi(t)$. In this observable case, the buyer's purchasing behavior could be better represented by resorting to the auxiliary function $\psi(t)$, defined as

$$\psi(t) = \min\{\phi(s) : s \leq t\}.$$

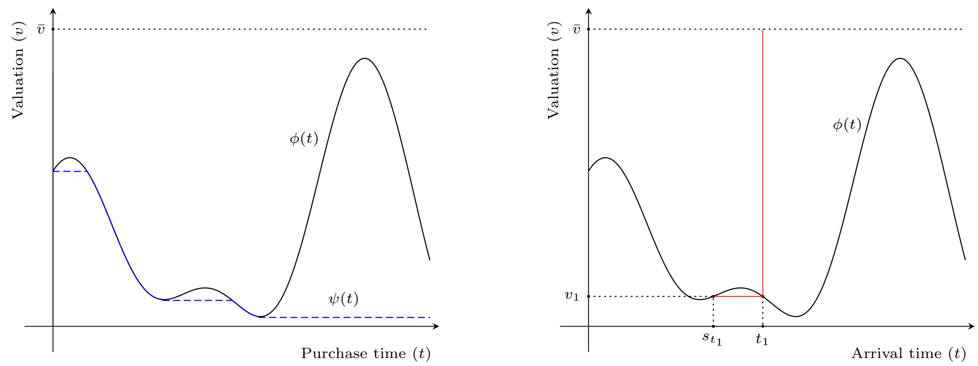
In other words, a customer arriving at time zero with valuation $\psi(t)$ will buy at time t. Due to the lower semi-continuity of p we have that ϕ is also lower semi-continuous and therefore, ψ is well defined (see Proposition 10 in the Appendix A.1 for a proof). The purchasing function $\psi(t)$ is the unique non increasing function that supports $\phi(t)$ from below (see Figure 1.1(a)). The instantaneous probability of selling at time t is given by $\mathrm{d}(1 - F(\psi(t)))$. With this observation we may write the seller's problem conditioned on the event that the buyer arrives at time 0:

$$
\begin{aligned}
[SPO_0] \qquad & \max_{p,\psi} \quad p(0)(1 - F(\psi(0))) + \int_0^\infty \mathrm{e}^{-\delta t} p(t) \ \mathrm{d}(1 - F(\psi(t))). \\
& \text{s.t.} \qquad t \in \arg\max_{s \geq 0} \ \ U(s, \psi(t)) \ \ \text{for all } t \geq 0.
\end{aligned}
$$

The first term in the objective function stands for the event where the customer buys immediately at time 0, and the second term accounts for his forward looking behavior. Following the standard assumption in the literature, we assume that the seller is more patient than the buyer and hence her discount factor δ verifies $\delta < \mu$. The incentive compatible constraint specifies that a consumer arriving at time zero with valuation $\psi(t)$ maximizes his utility at time $s = t$.

Note that every p feasible solution of the problem $[SPO_0]$ must be non increasing. Otherwise, there would exist $t > s > 0$ such that $p(t) > p(s) > 0$. Thus, $\psi(t) - p(s) > \psi(t) - p(t)$ and $\mathrm{e}^{-\delta s} > \mathrm{e}^{-\delta t} > 0$, and therefore, $U(s, \psi(t)) > U(t, \psi(t))$, which contradicts the definition of t in the constraint of $[SPO_0]$.

We can now extend the seller's revenue optimization problem to the case when the buyer arrives at time $\tau > 0$. Let R_τ be the seller's maximum expected revenue conditioned on the event that the buyer arrives at time τ. This corresponds to shifting the seller's revenue



(a) Observable case. Definition of the function $\psi(t)$. For a given function $\phi(t)$, a customer with valuation $\psi(t)$ arriving at $\tau = 0$ will buy at time t.

(b) Unobservable case. Characterization of a buyer purchasing at time t_1 including the one arriving exactly at t_1 with valuation $v \geq v_1$, and those arriving between s_{t_1} and t_1 with valuation v_1.

Figure 1.1: Consumer purchasing behavior

from $\tau = 0$ to $\tau > 0$, i.e., $R_\tau = \mathrm{e}^{-\delta\tau} R_0$, with R_0 being the objective function value of problem $[SPO_0]$. Finally, the ex-ante maximum expected revenue of the seller can be written as $R = R_0 \int_0^\infty \mathrm{e}^{-\delta\tau} g(\tau)\mathrm{d}\tau$, so that our assumption above on writing the seller's problem when the customer arrives at time zero is without loss of generality in terms of characterizing the seller's optimal pricing policy.

The formulation $[SPO_0]$ and the related R allows us to make a clear connection to the two alternative interpretations of our model: (i) there is a continuum of buyers with mass $\mathrm{d}(1 - F(\psi(t)))$ who buy at time t, for a total mass of 1 over the infinite horizon; and (ii) the seller is able to keep track of each of these buyers and post a personalized price curve $p(t)$.

Unobservable arrival case.

When the seller does not observe the buyer's arrival time, the price function that she has to set can only depend on the arrival time distribution G.

Although it is possible to formulate the seller's problem without any assumption over the threshold function ϕ, it is necessary to be careful on how to express her expected revenue when ϕ is not continuous. Thus, just for simplicity and because it does not affect the analysis in what follows, we describe the seller's problem under the assumption of ϕ being continuous.

Defining the point of time s_t as the last time previous to t where ϕ takes the same value as $\phi(t)$ (or $s_t = 0$ if such time does not exist, see Figure 1.1(b)), i.e.,

$$s_t = \sup\{l < t : \phi(l) = \phi(t)\} \vee 0,$$

the seller's problem can be described as follows:

$$[SPN] \quad \max_{p,\phi} \quad \int_0^\infty \mathrm{e}^{-\delta t} p(t) \left[(1 - F(\phi(t)))g(t) + 1_{\{\phi'(t) \le 0\}} (G(t) - G(s_t))(1 - F(\phi(t)))' \right] \mathrm{d}t.$$
$$\text{s.t.} \qquad t \in \arg\max_{s \ge t} \quad U(s, \phi(t)) \quad \text{for all } t.$$

The term in brackets stands for the probability of purchasing at time t. Within it, the first term $(1 - F(\phi(t)))g(t)$ represents the probability of arriving at time t with valuation $v \ge \phi(t)$ and hence purchasing immediately. This corresponds to the points in the vertical line in Figure 1.1(b); that is, we are accounting for a customer arriving in t_1 with valuation $v \ge v_1$.

The second term, $(G(t) - G(s_t))(1 - F(\phi(t)))'$, is the probability of purchasing at time t when arriving at any time between s_t and t with valuation $\phi(t)$, that is, the probability of being in the line connecting $\phi(s_{t_1})$ and $\phi(t_1)$ in Figure 1.1(b). Note that if the buyer has arrived before t and is still present at t, he will not buy if ϕ is increasing at t, and thus the latter term only holds at points where ϕ is decreasing.

The description of this optimization problem is included for completeness, but strictly speaking we will not solve it in our forthcoming development, but rather we would focus in a feasible pricing policy that would allow us to bound the ratio between the revenue from $[SPO_0]$ and $[SPN]$.

We conclude this section by making the connection with the two alternative model interpretations described in Section 1.1: (i) a model in which there is a continuum of infinitesimal buyers with point mass $1_{\{\phi'(t) \le 0\}}(G(t) - G(s_t))(1 - F(\phi(t)))' + (1 - F(\phi(t)))g(t)$, who buy at time t (and who have arrived before or at t), and (ii) the model with no price discrimination since all buyers face the same price curve posted at time zero.

1.5 Analysis of the model with an observable arrival

Given the argument stated in Section 1.4.2, to analyze the observable case it is sufficient to focus on the solution of $[SPO_0]$, where the buyer arrives at time 0.

The problem $[SPO_0]$ is difficult to solve because of its equilibrium constraint. Our approach will be to formulate a relaxed version of the problem by computing the first order condition of the equilibrium constraint. Then, by applying the Euler-Lagrange equation we will show that any solution of the relaxed problem also solves $[SPO_0]$. Moreover, we provide a characterization of the optimal price function as a solution of an ordinary differential equation, which turns out to have a unique solution for a large set of valuation distributions, and furthermore, it can be solved explicitly for at least for F being a uniform distribution.

To begin with, consider the incentive compatible constraint in problem $[SPO_0]$. If $t^* > 0$ is in the interior of the feasible region, then it must satisfy the first order condition $h(t) = 0$, where $h(s) = U_s(s, \psi(t))$, or equivalently, $\psi(t) = p(t) - \frac{p'(t)}{\mu}$. Now, consider the relaxed formulation:

$$[SPO_0^r] \quad \max_{p,\psi} \quad \int_0^\infty e^{-\delta t} p(t)(-\psi'(t)) f(\psi(t)) \, dt + p(0)(1 - F(\psi(0)))$$

$$\text{s.t.} \qquad \psi(t) = p(t) - \frac{p'(t)}{\mu} \quad \forall t \geq 0.$$

The feasible region of this constrained problem is larger than the one of $[SPO_0]$ and therefore, the objective function value of $[SPO_0^r]$ provides an upper bound of $[SPO_0]$.

Note that the problem $[SPO_0^r]$ can be written as the following unconstrained maximization problem on the price function $p(t)$:

$$\max_p \int_0^\infty e^{-\delta t} p(t) \left(-p'(t) + \frac{p''(t)}{\mu}\right) f\left(p(t) - \frac{p'(t)}{\mu}\right) \, dt + p(0)\left(1 - F\left(p(0) - \frac{p'(0)}{\mu}\right)\right). \tag{1.1}$$

Letting the integrand function be $G(t, p(t), p'(t), p''(t))$ and the expected revenue at time zero be r_0, problem (1.1) is equivalent to:

$$\max_p \int_0^\infty G(t, p(t), p'(t), p''(t)) \, dt + r_0.$$

Focusing on the first term above, the associated Euler-Lagrange equation that must be satisfied by an optimal price function $p(t)$ states that

$$\frac{d^2}{dt^2}\frac{\partial G}{\partial p''} - \frac{d}{dt}\frac{\partial G}{\partial p'} + \frac{\partial G}{\partial p} = 0.$$

Such function $p(t)$ is a stationary point of the functional

$$\int_0^\infty G(t, p(t), p'(t), p''(t)) \, dt.$$

After some algebra (detailed in Lemma 10 in the Appendix A.1) the Euler-Lagrange equation becomes:

$$f'(\psi(t))(\psi'(t))(-\delta p(t) + p'(t)) + f(\psi(t))[\delta(\delta - \mu)p(t) - 2\delta p'(t) + 2p''(t)] = 0, \tag{1.2}$$

where $\psi(t) = p(t) - \frac{p'(t)}{\mu}$. Of course, this equation can be written as a system of two first order differential equations by defining the auxiliary variable $u(t) = p'(t)$. Thus, by standard results on ODEs (see, e.g., Theorem 20.9 in [111]) we can show that there exists one and only one solution to the initial value problem given $p(0)$ and $p'(0)$ under mild continuity and differentiability conditions. These conditions hold if we for instance assume that $p(0) > 0$ and $p'(0) < 0$. While the former is natural to assume, the latter makes sense in the context of this observable case with price commitment, where a forward-looking consumer will never buy within an ε-interval starting at zero if the price is non decreasing at zero. Therefore, for a large set of valuation distributions, we have that the relaxed problem has exactly one solution.

Let us highlight that though we know that in the observable case $\psi(t)$ is non increasing by construction -and indeed we use this fact to formulate the seller's problem- $[SPO_0^r]$ could potentially have an optimal solution with a generic function $\psi(t)$. However, the following result establishes that this does not happen. In other words, if $\psi(t)$ corresponds to an optimal solution of the seller's relaxed problem, then it must be a non decreasing function. A proof is provided in Appendix A.2.1

Proposition 1. *Assume that the density function f is strictly positive. If the price function $p(t)$ is a continuously differentiable optimal solution of the relaxed problem $[SPO_0^r]$, then the optimal purchasing function $\psi(t) = p(t) - \frac{p'(t)}{\mu}$ is non increasing.*

Proposition 1, along with the upper bound defined by the solution to $[SPO_0^r]$, allow us to show that any solution of $[SPO_0^r]$ also solves the seller's problem $[SPO_0]$, proof provided in Appendix A.2.2.

Theorem 1 Any solution of the relaxed problem $[SPO_0^r]$ such that p is differentiable with continuous derivative also solves the seller's problem $[SPO_0]$.

Theorem 1 allows to simplify the solution of the seller's problem $[SPO_0]$. Furthermore, we show that the solution of the relaxed problem is a solution of an autonomous system of ordinary differential equations.

Thus, to solve the seller's problem $[SPO_0]$, first we formulate the Euler-Lagrange equation (1.2) and solve it. Its solution will depend on the initial values $p(0) > 0$ and $p'(0) < 0$. Then, we replace that solution in problem (1.1) and solve it in terms of the scalar variables $p(0)$ and $p'(0)$. Finally, using these optimal initial values, we can recover the optimal price function $p(t)$ and purchasing function $\psi(t)$ which are the optimal solutions of the original seller's problem $[SPO_0]$.

To conclude this subsection, we present a following technical result that states that if for a given parameter $c \in (0, 1)$, we need to ensure that the seller earns a fraction $1 - c$ of her expected revenue in problem $[SPO_0]$, it is enough to look at the problem until time $T = \ln(1/c)/\delta$. A proof of the Lemma is provided in Appendix A.2.3.

Lemma 1. *For a given parameter $c \in (0, 1)$, up to time $T = \ln(1/c)/\delta$, the seller's expected* *revenue in the observable arrival case is at least* $(1-c)R_0$*; i.e.,*

$$\int_0^T e^{-\delta t} p(t) d(1 - F(\psi(t))) \geq (1-c)R_0,$$

where $p(t)$ *is the solution from* (1.2) *to the observable case problem.*

For instance, if we want to reach at least half of R_0 and we normalize the seller's rate discount to 1, from this result we conclude that it is enough to consider the problem until $T = \ln(2)$. This implies that the time needed to get a big fraction of R_0 is relatively small and, moreover, it does not depend on the valuation distribution.

Before moving on to the unobservable case, in the next section we will do the analysis of the observable case for a particular instance of the problem. In particular, we will assume the valuation distribution is uniformly distributed between 0 and 1.

1.5.1 Uniform valuation case

Assume that the buyer's valuation is Unif$[0,1]$. Then, the problem $[SPO_0^r]$ becomes:

$$\max_p \int_0^\infty G(t, p(t), p'(t), p''(t)) dt + p(0)(1 - \psi(0)), \tag{1.3}$$

where $G(t, p(t), p'(t), p''(t)) = e^{-\delta t} p(t) \left(-p'(t) + \frac{p''(t)}{\mu}\right)$.

Formulating the Euler-Lagrange equation (1.2) in this case, we obtain

$$p''(t) - \delta p'(t) + \frac{\delta^2 - \delta\mu}{2} p(t) = 0,$$

which is a second order ordinary differential equation in the function $p(t)$ with constant coefficients and thus, it can be solved explicitly. In fact, its solution is given by:

$$p(t) = c_1 e^{\frac{1}{2}t(\delta - \sqrt{-\delta(\delta - 2\mu)})} + c_2 e^{\frac{1}{2}t(\delta + \sqrt{-\delta(\delta - 2\mu)})},$$

where c_1, c_2 are constants to be determined.

Note that $\delta + \sqrt{-\delta(\delta - 2\mu)} > 0$ and $\delta - \sqrt{-\delta(\delta - 2\mu)} < 0$ due to $\mu > \delta$. Therefore, the optimal pricing function is a sum of a negative exponential function and a positive exponential function. Thus, $p(t)$ could in principle go to infinity when t goes to infinity. However, $p(t) \in [0,1]$ for all t, and therefore, it must be the case that $c_2 = 0$. Thus, the optimal price function is a negative exponential function of the form:

$$p(t) = p(0) e^{\frac{1}{2}t(\delta - \sqrt{-\delta(\delta - 2\mu)})}.$$

In order to simplify the notation, define the positive constant $A = -\delta + \sqrt{-\delta(\delta - 2\mu)}$. We are left with computing $c_1 = p(0)$. Replacing the function $p(t)$ in the unconstrained problem (1.1), we can rewrite it as a maximization problem over $p(0)$ as follows:

$$\max_{p(0)} p^2(0) \int_0^\infty \mathrm{e}^{-(\delta+A)t} \left(\frac{A}{2} + \frac{A^2}{4\mu} \right) \mathrm{d}t + p(0) \left(1 - p(0) - p(0)\frac{A}{2\mu} \right).$$

Solving this problem, we obtain $p(0) = \frac{2\mu(\delta+A)}{(A+2\mu)(A+2\delta)}$. Noting that $p'(t) = -\frac{1}{2}Ap(0)\mathrm{e}^{-\frac{1}{2}At}$, we also obtain $p'(0) = -\frac{A\mu(\delta+A)}{(A+2\mu)(A+2\delta)}$.

Therefore, the pricing function that solves the Euler-Lagrange equation is given by

$$p(t) = \frac{2\mu(\delta + A)}{(A + 2\mu)(A + 2\delta)} \mathrm{e}^{-\frac{A}{2}t},$$

with corresponding purchasing function

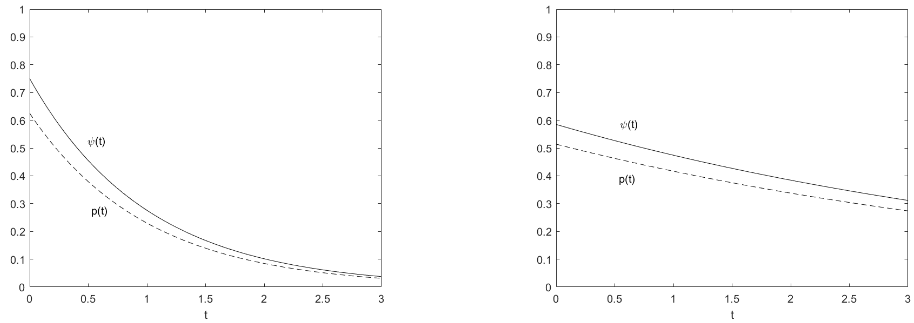
$$\psi(t) = \frac{\delta + A}{A + 2\delta} \mathrm{e}^{-\frac{A}{2}t}.$$

In conclusion, we have that the negative exponential functions $p(t)$ and $\psi(t)$ are the solutions of the seller's problem $[SPO_0]$ for a consumer's valuation Unif$[0, 1]$, and moreover, we have that the purchasing function is a positive multiplicative scaling the pricing function.

In what follows, we analyze the optimal curves obtained for some specific values of the discount rates μ and δ, corresponding to different levels of asymmetry in the patience of the seller and the buyer. Without loss of generality, we normalize the discount rate of the seller by setting $\delta = 1$.

In Figure 1.2, the left panel captures the case where the buyer is five times more impatient than the seller, whereas the right panel illustrates the scenario where he is only 50% more impatient. In panel (a), when the buyer is noticeably more impatient, we can observe that the optimal initial values of $p(0)$ and $p'(0)$ are greater than in panel (b), and that both price and purchasing optimal functions decrease faster. These curves reflect the fact that when facing a more impatient consumer (panel (a)), the seller will price more aggressively early in the horizon but will also drop the price quicker. Noting that the decreasing price pattern plays the role of a valuation discovery mechanism, the wider span of the pricing in (a) attempts to keep in the market a low valuation consumer by offering an attractive enough price relatively soon. On the contrary, when the buyer is more patient (panel (b)), the seller can offer a slow decaying price curve so that a consumer with mid to low valuation will buy later (compared to (a)) but at a higher price.

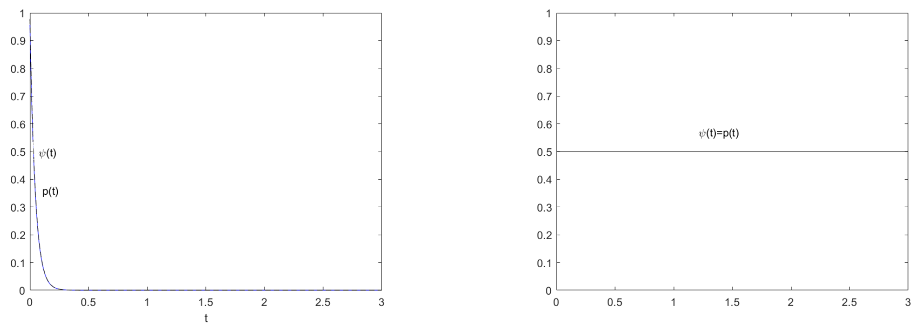
The fact that the seller takes advantage of the buyer's impatience is confirmed when computing the ex-ante expected revenue by solving $[SPO_0]$ in both cases, leading to values 0.3125 and 0.2574, respectively.



(a) The buyer is very impatient. Case $\delta = 1, \mu = 5$

(b) The buyer is slightly impatient. Case $\delta = 1, \mu = 1.5$

Figure 1.2: Optimal purchasing and price functions for different levels of asymmetry in the patience of the seller and the buyer.



(a) The buyer is extremely impatient. Case $\delta = 1, \mu = 1000$

(b) The buyer is very patient. Case $\delta = 1,\ \mu \to \delta$

Figure 1.3: Optimal purchasing and price functions for limiting asymmetries in the patience level of the seller and the buyer.

Figure 1.3 illustrates two limit scenarios for a normalized seller's discount rate $\delta = 1$. In panel (a) we consider the case in which the buyer is extremely impatient (with $\mu = 1000$). Here, the seller drops the price very quickly from 1 to 0, charging almost instantaneously the valuation of the buyer and extracting his whole surplus.

In Figure 1.3(b) we present the case in which the buyer's rate tends to 1. The optimal price and purchasing functions are the same and equal to 0.5 throughout the selling horizon. In this case, we recover the optimal auction of Myerson [110], with reservation price 0.5 and the buyer purchasing at time zero if and only if his valuation is at least 0.5. In this case, he pays the reservation price for the item.

The seller's advantage revenue-wise is even more emphasized, with values 0.4016 and 0.25, respectively.

1.6 Analysis of the model with unobservable arrival

Consider now the problem stated in Section 1.4.2 where the seller is not able to observe the arrival time of the buyer. Different from the previous observable case, where the seller knows the arrival time τ of the buyer and sets the price function $p(t)$ over the horizon $[\tau, \infty)$— even though, as explained before, the analysis was conducted without loss of generality by assuming $\tau = 0$—, in this case she commits to a price function at time zero.

This problem turns out to be very difficult in the general case. To partially overcome, we will focus on analyzing the seller's problem under a feasible pricing policy, with the objective of bounding the *value of observability*; that is, the ratio between the expected revenues under the observable case $[SPO_\tau]$ and the unobservable case $[SPN]$.

Our main result states that under a general valuation distribution, the *value of observability* is upper bounded by 4.911. However, in the case where the valuation distribution is monotone hazard rate, the bound is improved to $\mathrm{e} \approx 2.718$. To ease the exposition we first prove the latter result.

1.6.1 Monotone hazard rate valuation distribution

The case of monotone hazard rate valuation distribution turns out to be quite simple. We start this section by reviewing some basic concepts on the theory of optimal auctions introduced in the seminal work of Myerson [110]. Recall that the virtual valuation of the random variable $v \sim F$ is given by

$$J(v) := v - \frac{1 - F(v)}{f(v)} = v - \frac{1}{\rho(v)},$$

where $\rho(v) = f(v)/(1 - F(v))$ is the hazard rate function associated with the distribution F. The value $J(v)$ represents the expected value of the revenue that the seller may intend to collect from a bidder with valuation v, which naturally verifies $v > J(v)$. Alternatively, when considering the static price optimization problem of a seller trying to maximize the revenue function $r(p) = p(1 - F(p))$, the first order condition states that $J(p) = 0$. In other words, $J(p)$ stands for the marginal revenue function. As a consequence, an optimal monopoly reserve price p^* is defined as $p^* = J^{-1}(0)$.[5]

In what follows, we assume that the buyer's valuation is distributed according to a monotone (increasing) hazard rate distribution F and prove that the value of observability is upper bounded by e. Moreover, this bound is tight.

[5] More generally, the optimal reserve price is defined as $p^* = \max\{v : J(v) = 0\}$, and by convention, $p^* = \infty$ if $J(v) < 0$ for all v.

Indeed, we know from Section 1.4.2 that the optimal seller's expected revenue in the observable case is given by $R = R_0 \int_0^\infty \mathrm{e}^{-\delta t} g(t)\mathrm{d}t$, where R_0 is the objective function value of problem $[SPO_0]$ and therefore verifies $R_0 \leq \mathbb{E}(v)$, the expected value of the valuation drawn from F. Hence, the seller's expected revenue in the observable case is upper bounded by

$$\mathbb{E}(v) \int_0^\infty \mathrm{e}^{-\delta t} g(t)\mathrm{d}t.$$

For the unobservable case, consider the feasible, fixed pricing policy $p(t) = p^*$ for all t, where $p^* = J^{-1}(0)$ is the optimal monopoly price. Then, the seller's expected revenue is at least

$$\int_0^\infty \mathrm{e}^{-\delta t} p^*(1 - F(p^*))g(t)\mathrm{d}t = p^*(1 - F(p^*)) \int_0^\infty \mathrm{e}^{-\delta t} g(t)\mathrm{d}t.$$

Finally, by Lemma 3.10 (p.325) of Dhangwatnotai et al. [50], it follows that $p^*(1-F(p^*)) \geq \frac{1}{\mathrm{e}}\mathbb{E}(v)$, and this the claimed bound follows.

The bound is tight in the case of exponentially distributed valuation ($F(v) = 1 - \mathrm{e}^{-v}$) and a myopic buyer (with $\mu = \infty$), and when the seller does not discount revenues (i.e., $\delta = 0$). In this setting, in the observable case, the seller will announce a price curve that spans all the support $[0, \bar{v}]$ (e.g., $p(t) = 1/t$), and the consumer will buy immediately when his valuation $v = p(t)$. In this case, the ex-ante expected revenue is $\mathbb{E}(v) = 1$. In order to get the revenue for the unobservable case, the seller will offer a fixed price p maximizing $p(1 - F(p)) = p\mathrm{e}^{-p}$. This function is maximized at $p = 1$ with optimal revenue e^{-1}.

1.6.2 A feasible periodic price function

We start by noting that using fixed pricing does not work in general. For instance, consider the game where the buyer's valuation is distributed according to a truncated Pareto distribution with parameter 1, that is, with cdf $F(x) = (1 - 1/x)M/(M - 1)$ for $x \in [1, M]$, and again $\mu = \infty$ and $\delta = 0$. Here, we have that the expected value of the buyer's valuation is $M \ln M/(M - 1)$ whereas $p^*(1 - F(p^*)) = M/(M - 1)$, leading to the ratio $\mathbb{E}(v)/p^*(1-F(p^*)) = \ln M$ growing with M. Note then that the ratio grows arbitrarily large independent on the arrival distribution.

Thus, to bound the value of observability in the general case we need to consider a pricing policy that allows us to compare the expected revenue in the observable and unobservable case. We define it in Section 1.6.2 and present our main result in Section 1.6.3.

The feasible pricing policy $\hat{p}$ we consider is periodic and depends on the optimal pricing policy p of $[SPO_0]$. The length of the period will be $2T$ where T is such that until time T the seller's expected revenue in the observable case when the buyer arrives at time zero is

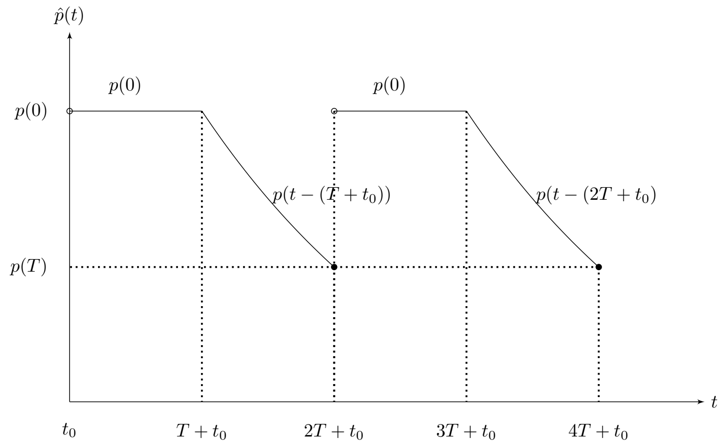


Figure 1.4: Periodic pricing policy $\hat{p}$ after performing a random shift and setting the origin at time t_0

big enough. In particular, the price function we use to bound the seller's expected revenue in the unobservable case is defined by

$$\hat{p}(t) = \begin{cases} p(0) & \text{if } t \in I_{2k-1},\ k \in \mathbb{N} \\ p(t-(2k-1)T) & \text{if } t \in I_{2k},\ k \in \mathbb{N} \end{cases} \tag{1.4}$$

where $I_{2k-1} = (2(k-1)T, (2k-1)T]$ and $I_{2k} = ((2k-1)T, 2kT]$ for $k \in \mathbb{N}$, and where the constant price $p(0)$ comes from the solution of $[SPO_0]$. Note that the function $\hat{p}$ is continuous at the points kT, for odd values of k.

Figure 1.4 shows the structure of the periodic pricing policy we will consider along the rest of the section, with origin at a value $t_0 \geq 0$. The fact of having a time origin set at t_0 is justified as follows. One element that makes the unobservable arrival case particularly challenging to analyze from a revenue computation perspective is the presence of the density $g(t)$ in the formulation $[SPN]$. In order to perform the analysis independently of the specific function g, let us first observe that by doing a random shift on the price function we can assume without loss of generality that the buyer's arrival time is uniformly distributed within a period of length $2T$.

More formally, suppose that we have a periodic function h with period $2T$ and consider a random shift, that is, for a random variable $t_0 \sim \text{Unif}[0, 2T]$, consider the function $\hat{h}(t) = h(t+t_0)$. Then, given that the buyer arrives in the interval $I_{2k-1} \cup I_{2k}$ of length $2T$, for some $k \in \mathbb{N}$, and denoting by X the random variable *arrival time*, we have the following:

$$\begin{aligned} \mathbb{P}(X \leq t | X \in (I_{2k-1} \cup I_{2k})) &= \mathbb{P}(X \leq t | X \in (2(k-1)T - t_0, 2kT - t_0]) \\ &= \mathbb{P}(X \in (2(k-1)T - t_0, t]). \end{aligned}$$

Letting s be the length of the interval $[2(k-1)T - t_0, t]$, i.e., $s = t - (2(k-1)T - t_0)$, the expression above verifies

$$\begin{aligned}\mathbb{P}(X \leq t | X \in (2(k-1)T - t_0, 2kT - t_0]) &= \mathbb{P}(X \in (2(k-1)T - t_0, 2(k-1)T - t_0 + s]) \\ &= \mathbb{P}(2(k-1)T - X < t_0 \leq 2(k-1)T - X + s) \\ &= \frac{t - (2(k-1)T - t_0)}{2T},\end{aligned}$$

which proves that X is uniformly distributed in $I_{2k-1} \cup I_{2k}$. Therefore, by applying a random shift over the function p to obtain $\hat{p}$, we can assume that buyer's arrival, conditional on the arrival interval, is Unif$[0, 2T]$, and that the function's new origin is t_0; that is, t_0 is the starting point of a period of length $2T$.

1.6.3 Revenue analysis

Along the rest of the chapter we will relabel the intervals of the function $\hat{p}$ and denote by $\tilde{I}_{2k-1}$ the range where $\hat{p}$ is constant, and will denote by $\tilde{I}_{2k}$ the range where $\hat{p}$ is a translation of the function p after performing the random shift.

We start by providing a simple lower bound for the seller's revenue within a limited time frame in the unobservable arrival case. Its proof is provided in Appendix A.3.1.

Lemma 2. *If the buyer is present at time τ being the beginning of a period $\tilde{I}_{2k}$ for some $k \in \mathbb{N}$, and has valuation $v \geq p(T)$, then the seller's expected revenue by offering the price function $\hat{p}$ in the unobservable case is at least the expected revenue earned up to time $2kT + t_0$ in the observable case with arrival time $(2k-1)T + t_0$.*

1.7 Bounding the value of observability

The value of observability $VO(G, F, \delta, \mu)$ of an instance of the problem with arrival distribution G, valuation distribution F, and discount rates δ and μ for the seller and the buyer, respectively, is defined as the ratio between the revenues in the observable and the unobservable cases. Accordingly, the value of observability VO is defined as the supremum of instance-specific values: $VO = \sup_{G,F,\delta,\mu} VO(G, F, \delta, \mu)$. In what follows, we provide upper and lower bounds for this worst-case value.

1.7.1 An upper bound for the value of observability

We are now able to bound the value of observing arrivals by considering the particular pricing policy $\hat{p}$ in (1.4) to give a lower bound of the seller's expected revenue in the unobservable case. To do so, we will only consider the buyer with valuation $v \geq p(T)$ when he arrives in an interval where the price is constant. We then obtain our main result which states that the value of observing the arrival is at most roughly 4.91 and its proof is provided in Appendix A.4.1.

This bound can be written as a function of W_{-1}, the negative branch of the well known Lambert function[6].

Theorem 2 For any valuation distribution and arrival time distribution, the value of observability is at most $-\frac{2W_{-1}(-1/(2\sqrt{e}))+1}{\left(e^{W_{-1}(-1/(2\sqrt{e}))+1/2}-1\right)^2} \approx 4.911$.

It is worth noting that our result is robust in the sense that it holds independently of the arrival distribution. That is, even in the case where the buyer arrival is adversarial—the worst possible for the seller—, we prove that the seller's expected revenue in the observable case is at most 4.911 times the seller's expected revenue if she does not observe the buyer's arrival.

1.7.2 A lower bound for the value of observability

Unfortunately, it is not straightforward to obtain a lower bound for the value of observability. The difficulty stems from the complexity of solving the unobservable case, even numerically, as discussed in Section 1.6. In order to partially overcome this difficulty, we consider a particular problem instance that can be solved numerically via dynamic programming.

Suppose that the valuation of the buyer is distributed uniformly in $[0,1]$, and that he arrives at one of two possible times. Specifically, we assume that the buyer arrives either at time 0 with probability β, or at time T with probability $1-\beta$, for some predetermined value $T>0$. We define the threshold valuation α as the value so that if the buyer arrives at time 0 with valuation $v \geq \alpha$, then he would buy before time T. This implies that, conditioned on that by time T the seller has not sold the item, the buyer's valuation is the mixture of two uniforms: (i) a Unif$[0,\alpha]$ accounting for the mass of buyers who arrived at 0 and decided to wait for a good price to be offered after T, with weight β, and (ii) a Unif$[0,1]$ distribution accounting for the buyer arriving at time T, with weight $1-\beta$. Assume also that the seller's discount rate is normalized to $\delta=1$, and that the buyer discounts the future at rate μ.

The general approach would be to decouple the problem in two independent subproblems that occur sequentially over time, by resorting to dynamic programming. Assume for now that the value of α is given. Then, we could solve the whole problem by backward induction. If there are no purchasing deviations—in the sense that if the buyer arrives with valuation at least α, he will not buy after T—, we can solve these two subproblems separately and then link them through the threshold α occurring at time T. Furthermore, each subproblem, once we condition on the information available at times 0 and T, corresponds to the observable

[6]The Lambert W function is defined as the multivalued function that satisfies $z = W(z)\exp(W(z))$ for any complex number z. If x is real then for $1/e \leq x < 0$ there are two possible real values of $W(x)$. We denote the branch satisfying $-1 \leq W(x)$ by $W_0(x)$—namely, the *principal branch*—, and the branch satisfying $W(x) \leq -1$ by $W_{-1}(x)$- referred to as the *negative branch*.

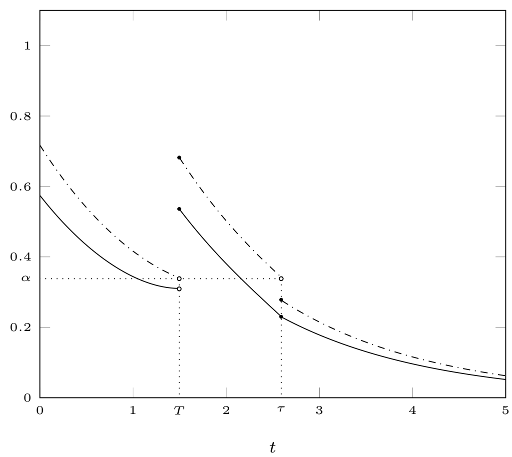
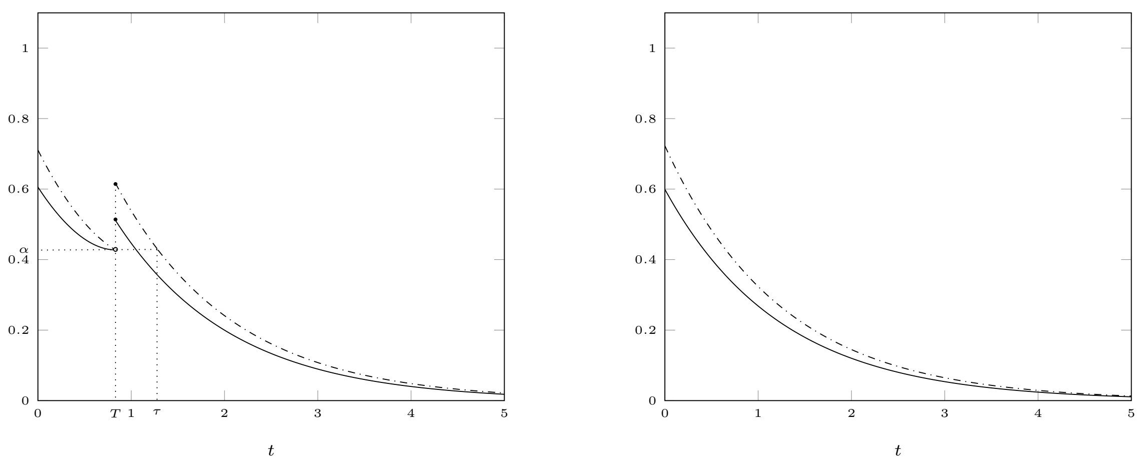


Figure 1.5: Price and purchasing functions for the special unobservable arrival case with Unif(0,1) valuation and two possible arrival times: 0 and T.

case (c.f. Section 1.5).

Starting from the second subproblem, defined over the interval $[T, \infty)$, we first guess the time τ by which the buyer would have purchased the item if and only if his valuation were above α, and then we can address the problem by solving two "observable" problems assuming a uniform valuation distribution. More specifically, the problem in $[\tau, \infty)$ is solved exactly as in Section 1.5.1; and the problem in the interval (T, τ) is solved similarly but fixing the value of the purchasing time function at time τ to be α. Finally, the problem in $[0, T]$ is solved with the same method, fixing the value of the threshold function at time T to be α. This whole procedure gives a price function and a purchasing time function that depend on α and τ—see Figure 1.5—, which are then optimized to maximize the seller's revenue. Note that as we have explicit solutions for the uniform valuation case (c.f. Section 1.5.1), the numerical part of the optimization to solve the unobservable case is only over these two parameters. Thus, putting all together, we can compute the value of observability for the instance defined by specific values of T, β, and μ. Finally, to obtain the best possible lower bound for this instance, we maximize the value obtained over these three parameters. This leads to a lower bound on the value of observability of 1.0173, which is attained by taking $T = 0.83$, $\beta = 0.67$ and $\mu = 3.9$. In particular, the expected revenue in the unobservable case is 0.2397 whereas the expected revenue in the observable case is 0.2439. Considering the instance that yields the lower bound, the left panel in Figure 1.6 illustrates the optimal pricing policy and threshold function for the unobservable case; whereas the right panel illustrates the functions for the observable case.



(a) Uobservable case. Optimal pricing policy and purchasing function for the instance with $T = 0.83$, $\beta = 0.67$ and $\mu = 3.9$.

(b) Observable case. Optimal pricing policy and purchasing function for the instance with $T = 0.83$, $\beta = 0.67$ and $\mu = 3.9$. Plot for arrival at time 0.

Figure 1.6: Optimal pricing policy and purchasing function for the instance attaining value of observability 1.0173

Chapter 2

Connection between posted price mechanisms and prophet inequalities[1]

2.1 Introduction

In the last few years online sales (particularly in ebay) have been moving from an auction format, to a posted price format [55] and the basic reason for this trend switch seems to be that posted price mechanisms are much simpler than optimal auctions yet efficient enough. In addition, in recent years several companies have started to apply *personalized pricing* to sell their products. Under this policy, companies set different prices (or offer different discounts) for different consumers based on purchase history or other factors that may affect their willingness to pay. Examples of such companies include Lexis-Nexis, Orbitz, and Safeway (see, e.g.,[123, 106, 90]).

The basic setting is as follows. Suppose a seller has an item to sell through a (sequential) posted price mechanism (PPM). In such a mechanism, consumers arrive one at a time and the seller proposes to each consumer a take-it-or-leave-it offer. The first consumer accepting the offer pays that price and takes the item. These type of mechanisms are very flexible and adapt very well to different scenarios [41]. Furthermore, their simplicity and the fact that strategic behaviour vanishes make them very suitable for a number of applications. Of course PPMs are suboptimal and therefore the study of their approximation guarantees has been an extremely active area in the last decade.

The recent survey by Lucier [102] is an excellent starting point in the area, where many variants of PPMs are described. In particular these can be: Anonymous (the offered price is the same for all consumers); Static (the possibly different prices to offer do not evolve as the

[1] This chapter of the thesis is based on joint paper with Jose Correa, Patricio Foncea and Victor Verdugo [45]

mechanism progresses), Order-Oblivious (the order in which agents arrive can be chosen by an adaptive adversary). Furthermore, we may even use PPMs when selling multiple objects and even have constraints on the subsets of consumers that can be served. Typical side constraints include selling multiple copies of an item, say we can serve at most k consumers, or having matroid constraints, or even more general downward-closed family constraints.

In the last decade a lot of effort from the computer science community has been devoted to understand the approximation guarantees that can be obtained through PPMs, where the natural benckmark is that given by the optimal mechanism of Myerson [110]. In an influential paper Chawla et al. [41] establish an interesting connection between (revenue maximizing) PPMs and prophet inequalities, a problem arising in optimal stopping theory. Here a gambler is faced to a sequence of random variables and has to pick a stopping time so that the value he gets is as close as possible to the expectation of the maximum of all random variables, interpreted as what a prophet, who knows the realizations in advance, could get. They implicitly show that any prophet type inequality can be turned into a posted price mechanism with the same approximation guarantee. This is obtained by noting that a PPM for revenue maximization can be seen as (threshold) stopping rule for the gambler but on the virtual values. As a consequence, the follow up work in the field concentrated on prophet inequities and then applied the obtained results to sequential posted price mechanisms.

In this chapter we fill a gap in this line of research by proving the converse of the latter result, namely, that any posted price mechanism can be turned into a prophet type inequality with the same approximation guarantee. The core of the result is a way to go back from virtual values to arbitrary distributions which may find applications beyond the scope of this thesis. This result amounts not only to apply approximation guarantees from prophet inequalities to PPMs, but also to carry over the lower bounds. We observe that actually though our reduction we can improve the best known lower bound for sequential PPMs (in which the arrival order is either random or selected by the seller) in the single item case, the k-uniform matroid case, the general matroid case, and the general downward-closed family case.

2.2 Contributions

We summarize contributions of Chapter 1 below.

Firstly, we provide a simple proof of the reduction from prophet inequalities to posted price mechanism, already studied by Chawla et al. [41]. The basic observation for obtaining this reduction is that price-based revenue maximization can be seen as prophet inequalities on the virtual values.

Secondly, we introduce a key result, namely *Valuation Mapping Lemma*, stating that for any distribution F there is another distribution G whose virtual value distributes according to F. That is, if we consider the operator that picks an arbitrary probability distribution over the nonnegative reals and returns the distribution of the ironed virtual valuation function, the Valuation Mapping Lemma states that this operator is surjective (onto) over the space of distributions. Interestingly, the lemma gives an explicit construction so we can easily interpret the thresholds as prices for the sequential posted price problem. It is somewhat surprising that this basic result was missing from the auction theory literature and we believe it may prove useful in settings beyond PPMs.

Thirdly, we answer the natural question arising from the above mentioned result of Chawla et al. [41] of whether the converse of their result also holds. In other words, we answer the question: does the existence of a PPM with a certain approximation guarantee implies the existence of a (threshold) prophet inequality with the same approximation guarantee? The main result of this chapter is to answer this question on the positive. Specifically, we prove that if we take an instance of the sequential posted price problem and there is a PPM obtaining a revenue within a factor α of Myerson's optimal mechanism, then there is a threshold stopping rule that obtains a fraction α of what a prophet could obtain. The main difficulty on the proof of this result comes from taking an arbitrary distribution in the prophet inequality problem and map it back to a sequential posted price problem. Here is where the valuation mapping lemma, that holds for arbitrary distributions, comes into play giving a way to go from virtual valuations to arbitrary distributions.

A remarkable feature of the Valuation Mapping Lemma is that when mapping a distribution F into another distribution G whose virtual value distributes like F, we can actually obtain a regular G (i.e., monotone non-decreasing virtual value). Of course, in principle there are many G's that satisfy the statement of the lemma, however we can identify one explicitly with this appealing property. Together with the reduction theorems, this implies that the posted price problem can be reduced to a prophet inequality problem, which can in turn be reduced to a posted price problem with regular valuation distributions. Therefore designing PPMs in general is equally hard from an approximation perspective to designing PPMs when the underlying valuations are regular.

It is worth mentioning that our result is very robust to different settings: it works with different ordering such as random, adversarial, or best possible, as well when there are multiple items and constraints on the allowed allocation sets.

Finally, we translate all known upper and lower bounds from PPMs into prophet inequalities, and back. A particular case where this gives new results in the case of sequential posted price mechanisms (SPM, [41]). Here the seller has to choose an order of the buyers and a pick a price for each buyer so as to maximize revenue. The current best known lower bound for this problem was obtained a decade ago by Blumrosen and Holenstein [28] and evaluates to $\sqrt{\pi/2} \approx 1.253$. Moreover, this lower bound is the best known lower bound when the feasibility constraint is a k-uniform matroid, a general matroid, and even the intersection of two matroids. With our result we can improve the lower bound in these settings to 1.341 by using the lower bound for the i.i.d. prophet inequality designed by Hill and Kertz [81] over three decades ago. On the other hand, we can also use our framework to derive a new simple and purely probabilistic proof of the ex ante relaxation bound on the revenue of Myerson's optimal mechanism [41, 7]. It is worth mentioning that, although our results are presented in the context of single-parameter mechanism design, they also follow more generally for multi-parameter domains by considering the reduction introduced by Chawla et al. [41].

2.3 Related literature

The connection between prophet inequalities and mechanism design was initiated by Hajiaghayi et al. [79] and further developed by Chawla et al. [40]. The key idea behind this connection was to implement prophet inequality's algorithms using thresholds that could be identified with prices. In particular, Chawla et al. [40] considers a pricing auction in adversarial order and obtain an upper bound of 4 with respect to the optimal mechanism. This was later improved by Chawla et al. [41] who gave several new approximation factors for PPMs in multiparameter settings, for the case of adversarial order (*order-oblivious posted price*) and when the designer was allowed to choose the order (*sequential posted price*). In the former case, they gave tight approximations for uniform and partition matroids (a constant factor of 2), and lower bounds for general matroids, intersection of matroids and downward-closed systems. Studying the prophet inequality problem, Kleinberg and Weinberg [91] gave improved upper bounds of 2 for general matroids and $4p-2$ for intersection of p matroids, as well as a lower bound of $O(p)$ for this second setting, while Babaioff et al.[18] and Rubinstein [119] presented a lower bound in downward-closed systems of $\log(n)/2\log(\log(n))$. Rubinstein also found a new upper bound of $O(\log(r)\log(n))$ on downward-closed systems (where r is the size of the largest feasible set) and a lower bound of $O(n)$ if the system is not downward-closed. In the case of sequential posted prices, Yan [134] used the correlation gap to find upper bounds of $\frac{\mathrm{e}}{\mathrm{e}-1}$ for general matroids and $p+1$ for intersection of p-matroids.

During the last years several other variants of the problem have been studied, including combinatorial prophet inequalities (see, e.g., [37, 120]), combinatorial auctions (see, e.g., [7, 61]), and polymatroids constraints for auctions (see, e.g., [54]). Attention has also been payed to settings with limited information or prior-independent, where the designer must learn the distribution in order to run the mechanism (see, e.g., [50, 43, 16, 73, 109, 17, 36]). Very recently, Dütting et al. [52] developed a general framework to obtain prophet inequalities using the concept of *balanced prices*, which allowed them to match and improve previous upper bounds in combinatorial auctions with sub-additive valuations and matroid constraints.

In the more specific setting where only one item must be allocated, the sequential posted price problem can be very simply stated as follows. Consider $n > 1$ buyers with valuation distributions represented by non-negative, independent random variables $Y_1, \dots, Y_n$. If buyer i is offered a price above his value then he takes the item and the seller obtains a gain equal to the offered price, otherwise the buyer is discarded and the seller continues with buyer i+1.

The goal is thus to find prices $p_1, \ldots, p_n$ maximizing the expected revenue of the seller. For this problem, the performance of anonymous pricing (i.e., all prices have to be equal) was considered by Alaei et al. [8] and Dütting et al. [53]. Chawla et al. [41] obtain upper bounds of 2 and $\frac{\mathrm{e}}{\mathrm{e}-1}$ for order-oblivious and sequential posted price mechanism respectively. Later, Esfandiari et al. [59] studied the prophet secretary setting were the arrival order is random, and provided an approximation factor of $\frac{\mathrm{e}}{\mathrm{e}-1}$ and a lower bound of 1.33 when the algorithm's prices may depend on the number of buyers that have rejected but not on their valuation distributions. The same approximation factor was found by Correa et al. [44] in the random order case, but where the algorithm's prices are static and had to be decided beforehand only based on the underlying distributions. Furthermore in the case of i.i.d. distributions, they found a 1.34 approximation factor which by the work of [81] is also tight. Abolhassani et al. [2] considered a relaxation of the i.i.d. setting under a large market assumption and showed that in both adversarial and random order a 1.36 approximation is possible.

Prophet-Inequalities

For fixed $n > 1$, let $X_1, \ldots, X_n$ be non-negative, independent random variables and T_n their set of stopping rules. A classic result of Krengel and Sucheston, and Gairing [93, 94] asserts that $\mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq 2 \sup\{\mathbb{E}(X_t) : t \in T_n\}$, and that 2 is the best possible bound. The study of this type of inequalities, known as *prophet inequalities*, was initiated by Gilbert and Mosteller [70] and attracted a lot of attention in the eighties (see, e.g., [81, 89, 91, 122, 121]). In particular Samuel-Cahn [122] noted that rather than looking at the set of all stopping rules one can (quite naturally) only look at threshold stopping rules in which the decision to stop depends on wether the value of the currently observed random variable is above a certain threshold (and possibly on the rest of the history). Two decades later the problem was brought up to the attention of the computer science community by the work of Chawla et al. [41] who noted a close connection between PPMs and prophet inequalities. The basic observation of Chawla et al. is that a PPM reduces to a (threshold) prophet inequality but applied to the virtual values [110]. The work of Chawla et al. was influential in that most follow-up work is directly on the context of prophet inequalities and then applies the implied results to PPMs.

The connection has also been useful to improve old results form prophet inequalities. Indeed, an interesting special case which we mention since we will use it later on, occurs when the random variables are also identically distributed. Here, the constant 2 can be lowered. Indeed, Hill and Kertz [81] provided family of "bad" instances from which Kertz [89] proved the best possible bound one could expect is $\beta \approx 1.341$, the unique solution to $\int_0^1 \frac{1}{y(1-\ln(y))+(\beta-1)} \mathrm{d}y = 1$. Hill and Kertz also proved a bound of $1 + 1/(\mathrm{e}-1) \approx 1.582$ which was very recently improved by Abolhassan et al. [2] to 1.355 and finally Correa et al. [44] proved that β is a tight value.

Organization

In Section 2.4 we introduce formally the online selection problem and the auction problem. We also introduce necessary notation and some useful remarks and results about distributions and its virtual valuation. In Section 2.5 we show the reduction from prophet inequalities to posted price mechanisms, result that is stated formally in Theorem 4. We remark that this approach has been considered extensively in mechanism design, and we include it for completeness, by showing a simple proof. In Section 2.6 we show our main result— stated formally in Theorem 5—, that is, the reduction from posted price mechanisms to prophet inequalities. Finally, in Section 2.7 we show some consequences of Theorem 5, including the improved lower bound for SPMs in more detail.

2.4 Preliminaries

We start this section by giving a formal definition of an instance of the two problems we study throughout this chapter: the online selection problem and the auction problem.

Online Selection Problem. An instance of this problem corresponds to a tuple $(X, \mathcal{F}, \mathcal{T})$, where X is the ground set of n elements and each set in $\mathcal{T} \subseteq 2^X$ is called *feasible selection*. For each element $x \in X$ there is a random variable w_x, called *weight*, distributed according to F_x with support contained in $\mathbb{R}_+$ and finite expectation, and $\mathcal{F} = \{F_x : x \in X\}$. We assume them to be independent. The random variables are presented in an order $\sigma : [n] \to X$, and an algorithm for the problem has to decide whether to select or not an element of X when arrived. An algorithm is correct if it outputs a feasible selection.

An algorithm is an *α-approximation* if the expected weight of the output selection is at least $\alpha \cdot \mathbb{E}\left(\max_{A \in \mathcal{T}} \sum_{x \in A} w_x\right)$, that is, an α fraction of the expectation of the maximum weight over feasible selections. In the latter, the expectation is taken over $\mathcal{F}$ and the (possibly) algorithm internal randomness.

Multi-Item Mechanism Design. Consider a single seller who provides a set of n items given by $\mathcal{I}$. For each item $\mathrm{i} \in \mathcal{I}$, there exists a buyer having a random valuation v_{i} for that item. We denote by G_{i} the distribution of the valuation v_{i}, and we assume this to have a support contained in $\mathbb{R}_+$ and that v_{i} is integrable. We denote by $\mathcal{G} = \{G_{\mathrm{i}} : \mathrm{i} \in \mathcal{I}\}$ the set of valuation distributions. There exists a set of *feasibility constraints* for the seller, $\mathcal{T} \subseteq 2^{\mathcal{I}}$, and every set in $\mathcal{T}$ is called a *feasible allocation*. Therefore, an instance for this problem is given by a tuple $(\mathcal{I}, \mathcal{G}, \mathcal{T})$.

This setting is known to be the *single-parameter domain*. We assume the valuation distributions to be independent, an they are known by the seller. Buyers arrive in an arbitrary order described by $\sigma : [n] \to \mathcal{I}$.

2.4.1 Myerson's optimal mechanism

In his seminal work, Myerson [110] characterizes the mechanism maximizing the revenue for single-parameter domains. In order to analize the optimization problem, he introduces a quantity called *virtual valuation*, that allows to solve the problem in an equivalent and simpler maximization setting.

Definition 1. *Given a random variable v with distribution G and density g, the virtual valuation of v corresponds to the function $\phi_G(t) = t - (1 - G(t))/g(t)$. We say that G is regular if ϕ_G is monotone non-decreasing.*

In the regular case, the optimal mechanism computes the virtual valuation for each buyer and then it allocates to a subset of them maximizing its total virtual value. Recall that a mechanism is called incentive-compatible if each player has a weakly dominant strategy of truthful reporting.

Theorem 3 ([110]) If the distributions in $\mathcal{G}$ are regular, the expected revenue of any incentive-compatible single-parameter mechanism $\mathcal{M}$ is equal to its expected virtual surplus, given by $\mathbb{E}\left(\sum_{x\in M} \phi_x^+(v_x)\right)$, where M is the allocation provided by the mechanism, and $\phi_x^+ = \max\{0, \phi_x\}$.

In particular, when the distributions are regular, the optimal mechanism by Myerson is incentive-compatible and so it satisfies the above conditions in the theorem. When the distributions are not regular, Myerson considered an *ironed* virtual valuation for its analysis [110], denoted by $\bar{\phi}_G$ when the valuation distribution is G. More specifically, take $Q(\theta) = \theta G^{-1}(1-\theta)$ and let R be the concave hull of Q, namely, $R(\theta)$ is given by

$$\min_{x,\theta_1,\theta_2\in[0,1]} \left\{xQ(\theta_1) + (1-x)Q(\theta_2) : x\theta_1 + (1-x)\theta_2 = \theta\right\}.$$

The ironed virtual valuation is $\bar{\phi}_G(t) = R'(1 - G(t))$. It is worth mentioning that, when the valuation is regular, the ironed virtual valuation corresponds to the virtual valuation, $\bar{\phi}_G = \phi_G$. If the context is clear, we omit the subscript on the notation for the (ironed) virtual valuation.

2.4.2 Posted-price mechanisms

In this chapter, we will work with a particular set of mechanisms, called *posted-price mechanism* (PPM). It works as follows: once a buyer arrives, the seller offers a price in a take-it-or-leave-it fashion, that is, if the valuation of the buyer for the item is greater than or equal to the price, then the buyer accepts the offer and purchases the item; it rejects otherwise.

The posted price mechanism, $\mathcal{M}$, upon arrival of a buyer preferring item $\mathrm{i} \in \mathcal{I}$, computes a price p_{i}. This price is a function of the *history* at time t, which corresponds to the current allocation and to the set of buyers arrived up to this point and their valuations. For short, we call the history $\mathcal{H}_t = (\sigma_t, A_{t-1}, \mathcal{V}_{t-1})$, where σ_t is the order in which the buyers arrived up to time t, A_{t-1} denotes the current allocation and $\mathcal{V}_{t-1} = \{v_{\sigma(j)} : j \in \{1, \ldots, t-1\}\}$ is the set of valuation realizations for the buyers so far arrived. Note that buyer i gets the item with probability $(1 - G_{\mathrm{i}}(p_{\mathrm{i}})$ and pays p_{i}. Therefore, the *expected revenue* of the mechanism is just $\mathbb{E}(\sum_{\mathrm{i} \in \mathcal{I}} p_{\mathrm{i}}(1 - G_{\mathrm{i}}(p_{\mathrm{i}})))$.

We assume that the mechanism fixes a price equal to $+\infty$ when by adding the item preferred by the buyer yields to an infeasible allocation.

2.4.3 About probability distributions and the virtual valuation

We recall that F is a distribution if it is a right-continuous and non-decreasing function, with limit equal to zero in $-\infty$ and equal to one in $+\infty$.

In general, F is not invertible but we work with its *generalized inverse*, given by $F^{-1}(y) = \inf\{t \in \mathbb{R} : F(t) \geq y\}$.

We denote by $\omega_0(F) = \inf\{t \in \mathbb{R} : F(t) > 0\}$ and $\omega_1(F) = \sup\{t \in \mathbb{R} : F(t) < 1\}$, and we call the interval $(\omega_0(F), \omega_1(F))$ the support of F. If $\omega_0(F)$ or $\omega_1(F)$ are finite, we include them in the support.

Below we present two useful technical results (see Chapter 2 in [87] for more details). We provide a proof in Appendix B.1.1 and Appendix B.1.2, respectively.

Proposition 2. *Let X be a real-valued random variable with distribution function F and F^{-1} its generalized inverse. Then, $F^{-1}(u) \leq x$ if and only if $F(x) \geq u$.*

Proposition 3. *Let X be a real-valued random variable with distribution function F and let U be uniformly distributed in $[0, 1]$. Then, $F^{-1}(U)$ has distribution F.*

2.5 Reduction from prophets to pricing

In this section we provide a reduction from the problem of online selection to multi-item auction problem by providing a PPM that calls a threshold based algorithm for the former. The idea of constructing posted-price mechanisms from existing prophet inequalities has been exploited extensively the last decade from the work of Chawla et al. [41]. In fact, this reduction is between lines in their work, although is not written in a general context. For the sake of completeness, we include a simple proof of this reduction in this section. The key idea behind the reduction is to convert an auction instance $(\mathcal{I}, \mathcal{G}, \mathcal{T})$ into an online selection one $(\mathcal{I}, \mathcal{G}^{\phi}, \mathcal{T})$, where $\mathcal{G}^{\phi} = \{G^{\phi}_{\mathrm{i}} : \mathrm{i} \in \mathcal{I}\}$ and G^{ϕ} is the distribution of the virtual valuation $\phi^+(v)$ where $v \sim G$.

We map thresholds in an online selection problem to prices in a posted price mechanism by feeding an online selection algorithm with the virtual valuations of the multi-item auction instance. In what follows, we denote by G^ϕ the distribution of the virtual values $\phi^+(v)$ where $v \sim G$. Given an instance $(\mathcal{I}, \mathcal{G}, \mathcal{T})$ for the multi-item auction problem, we construct an instance for online selection by picking one element for each item. The set of feasible selections is given exactly by the feasibility constraints of the auction problem, and the set of weight distributions are those given by the virtual valuation distributions, namely, $\mathcal{G}^\phi = \{G_{\mathrm{i}}^\phi : \mathrm{i} \in \mathcal{I}\}$. The order in which the elements of the instance are presented to the algorithm is the same order σ in which the buyers arrive. Given the instances above, we go from thresholds in the online selection problem to prices in the PPM as follows: upon the arrival of a buyer i, we run the online selection algorithm for the corresponding instance and compute the threshold τ_{i}. Then, if the distribution G_{i}^ϕ is regular, we define the price p_{i} as the generalized inverse of the virtual value ϕ_{i}^+ evaluated at the obtained threshold τ_{i}, otherwise, the mechanism needs to randomize between two prices.

Below we provide a pseudocode implementation for this algorithm when valuation distributions are regular. We denote by Alg both the set of elements chosen by the online selection algorithm and the rule that defines its thresholds. On the other hand, Mech represents the allocation provided by the resulting posted price mechanism.

Algorithm 1 From thresholds to prices.

Require: $(\mathcal{I}, \mathcal{G}, \mathcal{T})$ of the multi-item auction problem.
1: Initialize $M_0 \leftarrow \emptyset$.
2: **for** $t = 1$ to n **do**
3: Let $\mathrm{i} = \sigma(t)$, and compute threshold $\tau_{\mathrm{i}} = \mathrm{Alg}(\mathcal{H}_{t-1}, \mathcal{G}^\phi, \mathrm{i})$.
4: **if** $v_{\mathrm{i}} \geq (\phi_{\mathrm{i}}^+)^{-1}(\tau_{\mathrm{i}})$ **then** select i, $M_t \leftarrow M_{t-1} \cup \{\mathrm{i}\}$.
5: **else** reject i, $M_t \leftarrow M_{t-1}$.
6: Return Mech $= M_n$.

Before going to the main theorem of the section, we introduce a lemma that is used along the reductions. Recall that we denote by $\bar{\phi}$ be the *ironed* virtual value function [110]. Its proof is provided in Appendix B.2.1.

Lemma 3. *Let v be a random variable with regular distribution G. Then, for any $\tau \geq 0$,*

$$\mathbb{E}(\phi(v) \mid \phi(v) \geq \tau) = G^{-1}(1-q), \text{ with } q = \mathbb{P}(\phi(v) \geq \tau).$$

If the distribution is non-regular, there exist $q_1, q_2, x \in [0,1]$ such that $xq_1 + (1-x)q_2 = q$ and

$$\mathbb{E}(\bar{\phi}(v) \mid \bar{\phi}(v) \geq \tau) = \frac{xq_1 G^{-1}(1-q_1) + (1-x)q_2 G^{-1}(1-q_2)}{q}.$$

Using Lemma 3 and the mechanism obtained from the algorithm described above, we can make a reduction from any threshold rule in a prophet inequality setting to a posted price guarantee, stated formally in the following theorem and proved in Appendix B.2.2.

Theorem 4 ([41]) Suppose there exists an online selection algorithm based on thresholds that is an α-approximation for $(\mathcal{I}, \mathcal{G}^\phi, \mathcal{T})$ presented in order σ. Then, there exists a posted-price mechanism that is an α-approximation for $(\mathcal{I}, \mathcal{G}, \mathcal{T})$ presented in order σ.

2.6 Reduction from pricing to prophets

2.6.1 Reduction overview

Consider an instance $(X, \mathcal{F}, \mathcal{T})$ for the optimal stopping problem, and suppose we have access to an PPM single-parameter mechanism $\mathcal{M}$ that provides a guarantee over the ground set X and feasibility constraints $\mathcal{T}$. If we were able to find valuation distributions $\mathcal{G} = \{G_x : x \in X\}$ such that $\phi^+_{G_x}(v_x)$ has distribution F_x, where v_x has distribution G_x, then we could feed the mechanism $\mathcal{M}$ with the instance $(X, \mathcal{G}, \mathcal{T})$ of a multi-item auction problem, using the same order σ in which the elements of the ground set X are provided in the optimal stopping problem. We obtain prices $\{p_x : x \in X\}$ for each element, and then we can decide whether to select or not an element by mapping back to the domain of the online selection problem.

In particular, since w_x has the same distribution as $\phi^+_{G_x}(v_x)$, and they are all independent, by Theorem 3, the revenue of the mechanism on the instance $(X, \mathcal{G}, \mathcal{T})$ is the same provided by the weights of the elements selected, and so our online stopping algorithm provides a prophet inequality that preserve the approximation given by the mechanism $\mathcal{M}$ in the multi-item auction instance.

The rest of the section is devoted to prove that we can always find such set of valuations (Lemma 4 in Section 2.6.2) and to prove that the approximation is preserved (Theorem 5 in Section 2.6.3).

2.6.2 Valuation Mapping Lemma

In this section we introduce the key lemma that allows us to mapping from a weight distribution F to a valuation distribution G with virtual valuation distributed according to F.

Formally, we prove the following lemma.

Lemma 4 (Valuation Mapping Lemma)**.** *Let w be an integrable random variable with distribution F. Then, there exists a distribution G, such that $\phi^+_G(v)$ is distributed according to F, where v is a random variable with distribution G.*

The proof of the lemma (Appendix B.3.3)is constructive and we provide an explicit expression for the distribution G: we define G to be the generalized inverse of the function H

defined by

$$H(q) = \frac{1}{1-q}\int_q^1 F^{-1}(y)\mathrm{d}y,$$

in the interval $[0,1)$, $H \equiv 0$ in $(-\infty, 0)$ and $H \equiv 1$ in $[1, +\infty)$. Observe that $H(0) = \mathbb{E}(w)$, where $w \sim F$, and therefore H might be discontinuous in 0. The proof of Lemma 4 follows from a chain of propositions where we study the function H and we provide some useful properties of G and its virtual valuation.

Proposition 4. *The function H is continuous in $(0,1)$. Furthermore, there exists $T \in [0,1]$ such that H is strictly increasing in the interval $[0,T)$, and it is constant in the interval $[T,1)$.*

A proof of the proposition is provided in Appendix B.3.1. In fact, we show that $T = 1$ if F is continuous by the left in $t = \omega_1(F)$. Otherwise, if F is discontinuous in $t = \omega_1(F)$ then $T < 1$. In particular, if $\omega_1(F) = +\infty$, T=1. This good behaviour of H in the interval $(0,1)$ translates to G, in the sense that H is invertible in the whole $(0,1)$ except when $T < 1$ and so G has a discontinuity at $t = \omega_1(F)$. In other words, G is also strictly increasing and continuous in $(H(0), \omega_1(F))$.

Note that if the support of F is compact, H is a distribution.

In the following proposition, proved in Appendix B.3.2, we include the properties of G and its virtual valuation. In particular, it states that G is a regular distribution and we characterize its virtual value.

Proposition 5. *Let G be defined as above, extended it in the natural way: 0 in $(-\infty, 0)$ and 1 in $[1, +\infty)$. Therefore, the following holds:*

1. *The function G is a distribution and its support is $[\mathbb{E}(w), \omega_1(F)]$, where w is random variable with distribution F.*

2. *For all t in the support of G, $\phi_G(t) = F^{-1}(G(t))$.*

3. *The virtual valuation ϕ_G is non-decreasing. In particular, ϕ_G is non-negative and therefore $\phi_G^+ = \phi_G$.*

Finally, we introduce one more concept in order to show a property that plays a key role in the analysis of our algorithm in Section 2.6.3. Given non-decreasing functions $\eta : [0,1] \to [0,1]$ and $\nu : [a,1] \to [0,1]$ for some $a \in [0,1]$, we say that ν is a *non-linear stretching* of η if there exists $\xi : [a,1] \to [0,1]$ strictly increasing and continuous in $[a,1)$ such that $\nu = \eta \circ \xi$ in $[a,1)$. This type of transformation preserve some properties of the graph of the function η. In particular, there is a one to one correspondence between the intervals where η and ν are constant.

Proposition 6. *If ν is constant in the interval $[c, r)$, then η is constant in the interval $[\xi(c), \xi(r))$.*

Proof. Suppose that η is not constant over the interval $[\xi(c), \xi(r))$, that is, there exists $s \in (\xi(c), \xi(r))$ such that $\eta(s) > \eta(\xi(r))$. Since ξ is strictly increasing and continuous in (c, r), there exists $z \in (c, r)$ such that $\xi(z) = s$. Therefore, $\nu(z) = \eta(\xi(z)) = \eta(s) > \eta(\xi(r)) = \nu(r)$, which contradicts the fact that ν is constant in the interval $[c, r)$. □

Observe that since G is strictly increasing and continuous in $[\mathbb{E}(w), \omega_1(F))$, it follows that ϕ_G is a non-linear stretching of F^{-1}. In other words, if ϕ_G is constant in an interval $[c, r)$, then F^{-1} is constant over $[G(c), G(r))$.

Examples

Let us compute H and G for some particular valuation distributions F of w.

Uniform distribution. Consider that the random variable w is distributed according to a uniform distribution between zero and one, that is, $w \sim \text{Unif}[0, 1]$. Let us compute H and then G, represented in Figure 2.1.

Recall that

$$H(y) = \begin{cases} 0 & \text{if } y \in (-\infty, 0), \\ \frac{1}{1-y} \int_y^1 F^{-1}(s) \mathrm{d}s & \text{if } y \in [0, 1), \\ 1 & \text{if } y \in [1, +\infty) \,. \end{cases}$$

Then, considering $w \sim \text{Unif}[0, 1]$ we have that

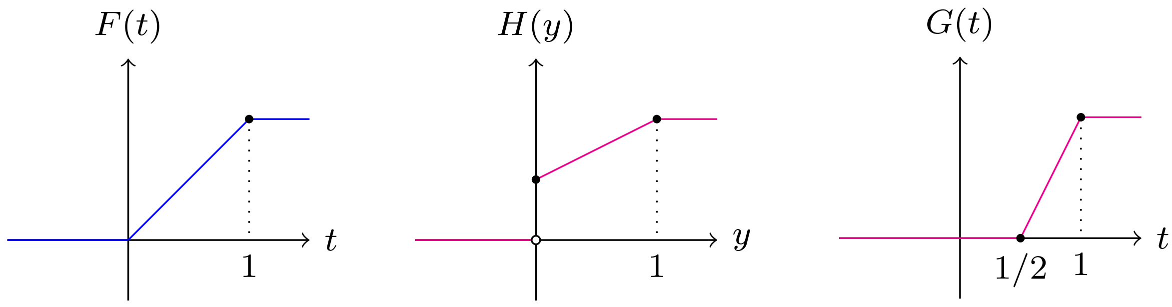
$$H(y) = \begin{cases} 0 & \text{if } y \in (-\infty, 0), \\ \frac{1}{2}(1 + y) & \text{if } y \in [0, 1), \\ 1 & \text{if } y \in [1, +\infty), \end{cases}$$

and computing its generalized inverse we obtain G and it is given by

$$G(t) = \begin{cases} 0 & \text{if } t \in (-\infty, 1/2), \\ 2t - 1 & \text{if } t \in [1/2, 1), \\ 1 & \text{if } t \in [1, +\infty). \end{cases}$$

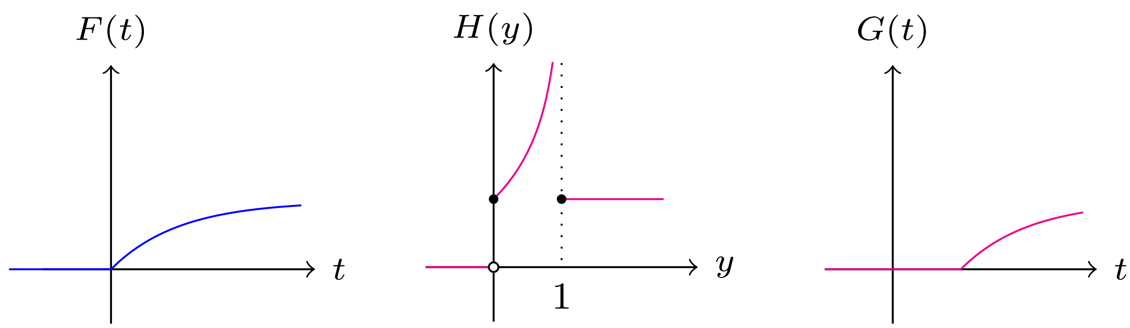
Therefore, if $w \sim \text{Unif}[0, 1]$, $v \sim G$ is also uniform but between $1/2$ and 1. Moreover, the virtual value $\phi_G(v)$ is distributed according to F, i.e., $\text{Unif}[0, 1]$.

Exponential distribution. We now consider a random variable without compact support. In



(a) F distribution (b) H distribution (c) G distribution

Figure 2.1: F,H and G functions for the uniform example



(a) F distribution (b) H function (c) G distribution

Figure 2.2: F, G and H functions for exponential example.

particular, we consider $w \sim \text{Exp}(1)$. Note that $F^{-1}(y) = \ln(1/1 - y)$ for $y \in [0, 1)$. Then,

$$H(y) = \begin{cases} 0 & \text{if } y \in (-\infty, 0), \\ 1 - \ln(1 - y) & \text{if } y \in [0, 1), \\ 1 & \text{if } y \in [1, +\infty). \end{cases}$$

Computing the generalized inverse of H and extended it naturally, we obtain the distribution G given by

$$G(t) = \begin{cases} 0 & \text{if } t \in (-\infty, 1], \\ 1 - \mathrm{e}^{1-t} & \text{if } t \in (1, \infty). \end{cases}$$

Note that the virtual value associated to $v \sim G$ has distribution F, as is stated in Lemma 4. In Figure 2.2 are F, G and H representations.

2.6.3 From posted prices to online selection

As it was mentioned in Section 2.6.1, given an instance $(X, \mathcal{F}, \mathcal{T})$ for the online selection problem, the idea is to feed a single-parameter mechanism, $\mathcal{M}$, by constructing a set of valuations $\mathcal{G}$ using the Valuation Mapping Lemma over $\mathcal{F}$. Specifically, for each $x \in X$, we apply Lemma 4 to F_x, obtaining a random variable (valuation) v_x distributed according to G_x and with $\phi_{G_x}(v_x)$ distributed according to F_x. To not overload notation, in what follows we call ϕ_x the virtual valuation of G_x. We denote by $\mathcal{G} = \{G_x : x \in X\}$ the set of valuation distributions and consider the instance for the PPM given by $(X, \mathcal{G}, \mathcal{T})$.

Upon the arrival of an element (buyer) $x \in X$, we obtain a price p_x given by the mechanism $\mathcal{M}$. We then define a threshold for an element to be selected, taking into account that ϕ_x is guaranteed to have distribution F_x.

The idea of the algorithm is as follows. For an element $x \in X$, if ϕ_x is strictly increasing in neighbourhood of p_x, then x is selected if and only if $w_x \geq \phi_x(p_x)$. Otherwise, ϕ_x is flat in an interval containing p_x and this translates into a randomized tie-breaking for determining the threshold stopping rule. More specifically, consider the *boundary prices* given by

$$p_x^- = \inf\{p \in \mathbb{R} : \phi_x(p) = \phi_x(p_x)\},$$
$$p_x^+ = \sup\{p \in \mathbb{R} : \phi_x(p) = \phi_x(p_x)\}.$$

In words, if p_x falls in an interval where ϕ_x is constant, p_x^- is the left-most price with the same virtual valuation, and p_x^+ is the right-most value of that interval. Observe that $\phi_x(p_x^+)$ is not necessarily equal to $\phi_x(p_x)$.

If $w_x > \phi_x(p_x)$, the algorithm selects x with probability 1. If $w_x = \phi_x(p_x)$, the algorithm selects x with probability

$$\theta_x = \frac{G_x(p_x^+) - G_x(p_x)}{G_x(p_x^+) - G_x(p_x^-)},$$

and rejects it with probability $1 - \theta_x$. In any other case the element x is rejected. We now provide the pseudocode of the algorithm for constructing the thresholds for the optimal stopping problem described above. We assume the elements arrive following the order σ. Then, we state formally the main theorem, which is proven in Appendix B.3.4.

Theorem 5 Let $(X, \mathcal{F}, \mathcal{T})$ be an instance of the online selection problem, and $(X, \mathcal{G}, \mathcal{T})$ the instance of the multi-item auction obtained by the Valuation Mapping Lemma. Suppose the mechanism $\mathcal{M}$ is an α-approximation for $(X, \mathcal{G}, \mathcal{T})$ presented in order σ. Then, Algorithm 2 is an α-approximation for $(X, \mathcal{F}, \mathcal{T})$ presented in order σ.

2.7 Implications of the reduction

A direct consequence of Theorem 5 and the Valuation Mapping Lemma is that we obtain lower bounds for the guarantees of PPM's by considering lower bound instances of the online selection problem. In particular, we improve the previous known lower bounds for SPM when constraints are on the form of downward closed families, from $\log n/(3 \log\log n)$ to

Algorithm 2 From posted prices to thresholds.

Require: $(X, \mathcal{F}, \mathcal{T})$ of the online selection problem.

Initialize $A_0 \leftarrow \emptyset$.
for $t = 1$ to n **do**
  Let $x = \sigma(t)$, and compute price $p_x = \mathcal{M}(\mathcal{H}_{t-1}, \mathcal{G}, x)$,
  if $p_x^- = p_x^+$ and $w_x \geq \phi_x(p_x)$ **then** select x, $A_t \leftarrow A_{t-1} \cup \{x\}$;
  else if $p_x^- < p_x^+$ **then**
    if $w_x > \phi_x(p_x)$ **then** select x, $A_t \leftarrow A_{t-1} \cup \{x\}$;
    else if $w_x = \phi_x(p_x)$ **then** select x with probability θ_x, $A_t \leftarrow A_{t-1} \cup \{x\}$;
      reject x with probability $1 - \theta_x$, $A_t \leftarrow A_{t-1}$.
  else reject x, $A_t \leftarrow A_{t-1}$.
Return Alg $= A_n$.

$\log n/(2 \log\log n)$ [41, 119], and in the k-uniform matroid setting from 1.253 to 1.341 [28, 81] (actually, this lower bound also applies to general matroid constraints or even intersection of matroids). Additionally, using the results from Göbel et al. [72], it is possible to derive a new lower bound for PPM where the feasibility is given by stable sets in graphs, of $\Omega(\log n/\log^2 \log n)$. Finally, we also provide an alternative proof for the ex-ante relaxation lemma first obtained by Chawla et al. [40].

2.7.1 Improved lower bounds for PPMs

Consider for instance the single item sequential posted price problem as defined by Chawla et al. [41]. Here we have one seller and n buyers with valuations given by independent random variables $Y_1, \ldots, Y_n$. In a sequential posted price mechanism (SPM), the seller has to choose an order of the buyers and pick a price for each buyer so as to maximize revenue. As usual, buyers simply decide to either buy at the offered price (if their valuation is above the price) or reject the offer and get nothing. Despite the significant amount of work in the area, the current best known lower bound for this problem was obtained a decade ago by Blumrosen and Holenstein [28]. The instance is quite simple; all buyers have i.i.d. valuations distributed according to $F(v) = 1 - 1/v^2$. Simple calculations show that the expected revenue of the optimal mechanism is $\Gamma(1/2)\sqrt{n}/2$, while that of the optimal SPM is $\sqrt{n/2}$. Thus, the ratio is $\Gamma(1/2)/\sqrt{2} = \sqrt{\pi/2} \approx 1.253$. Rather surprisingly, this lower bound is the best known lower bound when the feasibility constraint is a k-uniform matroid, a general matroid, and even the intersection of two matroids.

To see that the lower bound can be improved we consider the lower bound for the i.i.d. prophet inequality designed by Hill and Kertz [81] over three decades ago. Hill and Kertz considered the problem of finding the best constant a_n such that for n i.i.d. random variables the expected gambler's gains are within a factor a_n of that of the gambler. They were able to characterize a_n through a recursion and also to find the instance that exactly achieves this gap of a_n. In follow-up work, Kertz [89] proves that a_n converges to $\beta \approx 1.341$, the unique solution to the integral equation $\int_0^1 1/(y(1-\ln(y))+(\beta-1))\mathrm{d}y = 1$.

With the Valuation Mapping Lemma we map back the distributions of the instances of Hill and Kertz to distributions for the sequential posted price problem. Since the distributions used in Hill and Kertz's instances are i.i.d., those for the sequential posted price problem will also be i.i.d. and Theorem 5 guarantees that the gap of $\beta \approx 1.341$ will be preserved. It is important to note here that the instances of Hill and Kertz are discrete and therefore we need to use the full power of the valuation mapping lemma for irregular, discrete, and piecewise constant distributions. With the above discussion we conclude the following corollary.

Corollary 1. *For 1-uniform matroid feasibility constraints (also k-uniform, general matroid, or intersection of two matroids), there is no SPM with approximation guarantee better than* $\beta \approx 1.341$*, where* β *is the unique solution to the integral equation*

$$\int_0^1 \frac{1}{y(1-\ln(y))+(\beta-1)}\mathrm{d}y = 1.$$

There are other situations in which our main result gives improved lower bounds for PPMs. This includes the case in which the feasibility set corresponds to stable sets in a graph. Here the lower bounds of Göbel et al. [72] which, for example, state that even for interval graphs one cannot hope to obtain an online algorithm with a performance guarantee better than $\Omega(\log n/\log^2 \log n)$, carry over to PPMs.

2.7.2 Bounding the revenue of Myerson's optimal mechanism

An important tool that has been widely used in the design of PPMs (and beyond) is the so called Ex-ante relaxation which basically states that an upper bound of the revenue of Myerson's optimal mechanism is obtained by bringing the ex-post allocation constraints to ex-ante constraints which will be satisfied in expectation. In the single item setting this upper bound was obtained by Chawla et al. [41] and developed further by Alaei [7]. Here we provide a simple and purely probabilistic proof of this result using the tools developed in this chapter. Before formalizing the proof we need the following simple technical result (see e.g. [44, Corollary 2.2]).

Lemma 5. *Let* $Y_1, \dots, Y_n$ *be non-negative random variables distributed according to* $F_1, \dots, F_n$ *respectively. For each* $\mathrm{i} \in \{1, \dots, n\}$*, set* $q_\mathrm{i} = \mathbb{P}(Y_\mathrm{i} = \max_{j\in\{1,\dots,n\}} Y_j)$ *and let* α_i *be such that* $1 - F_\mathrm{i}(\alpha_\mathrm{i}) = q_\mathrm{i}$*. Then,*

$$\mathbb{E}\left(Y_\mathrm{i} \;\middle|\; Y_\mathrm{i} = \max_{j\in\{1,\dots,n\}} Y_j\right) \leq \mathbb{E}\left(Y_\mathrm{i} \mid Y_\mathrm{i} > \alpha_\mathrm{i}\right).$$

If $F_\mathrm{i}^{-1}(1-q_\mathrm{i})$ *is empty for some* $\mathrm{i} \in \{1, \dots, n\}$*, the result still holds via randomization.*

Our proof, which we show only for the unit demand case, is very simple and uses only probabilistic arguments. So consider a single-item auction instance, that is a set of buyers

$\mathcal{I}$, valuation distributions are $\mathcal{G} = \{G_{\mathrm{i}} : \mathrm{i} \in \mathcal{I}\}$ and the feasibility constraints are $\mathcal{T} = \{\emptyset\} \cup \{\{\mathrm{i}\} : \mathrm{i} \in \mathcal{I}\}$.

Lemma 6 ([41])**.** *If the valuation distributions in $\mathcal{G}$ are regular, then the expected value of Myerson's optimal auction is upper bounded by*

$$\sum_{\mathrm{i}\in\mathcal{I}} G_{\mathrm{i}}^{-1}(1-q_{\mathrm{i}})q_{\mathrm{i}}\,,$$

where q_{i} is the probability that the optimal auction assigns the item to buyer $\mathrm{i} \in \mathcal{I}$. Furthermore, for every $\mathrm{i} \in \mathcal{I}$ (with regular or non-regular distribution) there exist two prices $\underline{p_{\mathrm{i}}}$ and $\overline{p_{\mathrm{i}}}$, with corresponding probabilities $\underline{q_{\mathrm{i}}}$ and $\overline{q_{\mathrm{i}}}$, and a number $0 \le x_{\mathrm{i}} \le 1$, such that $x_{\mathrm{i}}\underline{q_{\mathrm{i}}} + (1-x_{\mathrm{i}})\overline{q_{\mathrm{i}}} = q_{\mathrm{i}}$, and the expected revenue of Myerson's optimal auction is bounded from above by

$$\sum_{\mathrm{i}\in\mathcal{I}} x_{\mathrm{i}}\underline{p_{\mathrm{i}}}\,\underline{q_{\mathrm{i}}} + (1-x_{\mathrm{i}})\overline{p_{\mathrm{i}}}\,\overline{q_{\mathrm{i}}}\,.$$

Proof. Recall that by Theorem 3 the expected revenue of optimal mechanism is given by the expectation of the maximum over the virtual valuations. By conditioning and using Lemma 5, we can express the revenue of the optimal auction as

$$\begin{aligned}
\mathbb{E}\left(\max_{\mathrm{i}\in\mathcal{I}} \phi_{\mathrm{i}}^{+}(v_{\mathrm{i}})\right) &= \sum_{\mathrm{i}\in\mathcal{I}} \mathbb{E}\left(\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) \mid \phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) = \max_{j\in\mathcal{I}} \phi_{j}^{+}(v_{j})\right) q_{\mathrm{i}} \\
&\le \sum_{\mathrm{i}\in\mathcal{I}} \mathbb{E}(\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) \mid \phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) > \alpha_{\mathrm{i}})q_{\mathrm{i}},
\end{aligned}$$

where α_{i} is a value for which $\mathbb{P}(\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) > \alpha_{\mathrm{i}}) = q_{\mathrm{i}}$. We recall Lemma 3 to conclude the upper bound on the optimal mechanism, and the existence of such prices. □

Chapter 3

Dynamic optimization problems: a unifying approach[1]

3.1 Introduction

Dynamic optimization problems with resources constraints arise across a variety of disparate applications. For example, retailers engage in dynamic pricing with inventory constraints, airlines and hotels engage in dynamic capacity allocation problems with limited seat or room capacity, advertisers engage in real-time bidding campaigns with limited budget. Due to the importance and centrality of these problems, various classes of dynamic optimization problems have received significant attention in industry but also across academic communities in Operations Research, Computer Science, and Economics. A significant focus of the literature has been on the development of efficient algorithms to optimize performance.

While the literature on these problems is rich and extensive (we discuss the literature in detail as we present our main results and associated corollaries), studies have focused on specific applications, or classes of applications. As such, arguments are specialized for specific settings and do not directly apply to other settings, typically requiring to re-develop, from scratch, proofs when faced with a new type of dynamic optimization problems with resource constraints. While, from a practical perspective, problems such as those mentioned above can appear very different, these problems do admit some common mathematical structure. In the present work, we elucidate such common structure and derive important theoretical implications of such commonalities.

[1] This chapter of the thesis is based on a working paper with Omar Besbes and Santiago Balseiro [19].

3.2 Contributions

We summarize contributions of Chapter 3 below.

Firstly, we introduce and define a general class of problems: *dynamic resource constrained reward collection* (DRCRC) problems. This class admits as special cases a variety of problems studied completely separately in the literature. Broadly speaking, a DRCRC problem is defined as follows. A decision-maker is endowed with some resources at time 0 and faces a finite (discrete) time horizon. At each period, the decision-maker is presented with a stochastic opportunity (independent of other periods), and must select an action; the action leads to some stochastic resource consumption and reward collection. The goal of the decision-maker is to select a sequence of actions to maximize her total expected rewards subject to the resource constraints. We assume that the decision-maker knows the distribution of the various stochastic components, and, as such, this problem can be formulated as a discrete and finite time dynamic program, with the state given by the vector of resources available.

Notably, the DRCRC class of problems generalizes and brings under the same umbrella a host of classical problems studied separately. In particular, we show how the proposed class of DRCRC problems encompass the following classical problems: *Network dynamic pricing* problems (see, e.g., [66]) *Dynamic bidding in repeated auctions with budgets*, (see, e.g., [20]), *Network revenue management* problems (see, e.g., [129]), *Choice-based revenue management* problems (see, e.g., [128]), *Stochastic Depletion* problems (see, e.g., [39]), *Order fulfillment* problems (see, e.g., [3]), and *Online matching* problems (see, e.g., [5]).

Secondly, we provide a unified analysis of a “fluid" certainty-equivalent control. Although from a theoretical perspective, DRCRC problems can be formulated through a dynamic program, one natural question is whether the DRCRC formulation lends itself to analysis (beyond a generic analysis of a general dynamic program) that can applied to all special cases, or whether problems should be specialized first to be able to derive properties of interest. We indeed show that the general DRCRC formulation can lead to unified analysis, through the study of a central heuristic in the stochastic dynamic optimization literature.

Thus, our second layer of contribution is the analysis domain. In particular, we characterize the performance of a classical “fluid" certainty-equivalent control for the general DRCRC class of problems.

In more detail, solving even a special case of a DRCRC problem to optimality is typically impossible due to the curse of dimensionality; indeed the state space is driven by the number of resources. This has brought forward the need for heuristics for such problems, and many such heuristics have been developed for subsets of the problems above. A notable heuristic for dynamic optimization problems is the so-called certainty-equivalent heuristic, which involves solving a deterministic problem in each period by using proxies for random quantities, implementing the prescribed decisions for that period, and repeating the process over time. Such certainty-equivalent heuristics have been shown to be near-optimal under some conditions in various special cases of DRCRC problems. A notable example of a certainty-equivalent heuristic is the so-called "fluid" one, in which the random quantities are replaced by their expectations. We will refer to this heuristic as CE. Such policies are sometimes also referred to as "re-solving" or "model predictive control" in the various related streams of literature. To analyze the performance of the heuristic, we measure the gap between the optimal performance and that of the CE heuristic, and characterize the dependence of this gap on the "scale" of the system.

More specifically, we establish two types of sufficient conditions for the CE heuristic to lead to a "small" performance gap. The first set of sufficient conditions are stated in terms of local smoothness of the initial deterministic proxy problem as a function of the resource vector. The second, alternative, set of sufficient conditions are expressed in terms of the dual Lagrangian function. In particular, the analysis leads to a dichotomy between two fundamental cases: that when the set of actions is a continuum, and that when it is finite. For the former, the CE heuristic guarantees $O(\log T)$ performance gap (where T is the length of the horizon), whereas for the latter, it guarantees $O(1)$ gap.

In essence, the analysis establishes that the classes of DRCRC problems, under said sufficient conditions is "easy" in that the CE heuristic is extremely effective. Intuitively, the CE heuristic enables the decision-maker to implement good decisions through the proxy problem while controlling very closely the path of the resource constraints.

Thirdly, after developing our theoretical results and the dichotomy between finite and continuum of actions, we establish the mapping from a series of classical problems to a DRCRC problem and state the corollaries that one obtains from the general analysis of the CE heuristic. This allows to recover versions of various existing results but also to obtain such results under weaker conditions (see, e.g., the case of dynamic pricing in Section 3.8.1), or to obtain new results in the literature for the performance of the fluid CE heuristic (see, e.g., the case of dynamic bidding with budgets in Section 3.9.2 or the case of dynamic assortment optimization in Section 3.8.2). We discuss the related literature in detail when we discuss the various specialized problems.

Overall, this paper introduces a novel general formulation of dynamic optimization problems. We illustrated how this formulation lends itself to analysis through a unified analysis of the CE heuristic. As such, the DRCRC class offers a "useful" and powerful intermediate class of problems between the specialized versions previously studied in the literature and a fully general dynamic program, and this work opens up the possibility of further generalizations of arguments developed for special cases of DRCRC problems.

3.3 Related literature

We left the description of the related literature to Section 3.8 and Section 3.9, where we establish the mapping from a series of classical problems to a DRCRC problem.

Organization

In Section 3.4 we describe the model for the class of DRCRC problems. In Section 3.5 we present the certainty equivalent heuristic (CE) we analyze along the work, whereas in Section 3.6 we study its performance for the general class of DRCRC problems, providing an explicit bound for the reward loss, formally stated in Theorem 6. In Section 3.7 we present sufficient conditions for the guarantee of CE, depending on whether the set of actions is finite or not. Finally, in Sections 3.8 and 3.9 we review a set of special cases of the proposed DRCRC class of problems. For each class, we first establish when and how it falls under a special case of a DRCRC problem. In turn, we establish implications of Theorem 6. As we will see, the result allows to recover some existing results in the literature as special cases, but also have as a direct corollary new results for classes of problems studied in the literature.

3.4 Model

We consider a dynamic decision-making problem with a finite time horizon T, over which a decision-maker collects rewards subject to resource constraints. We refer to this problem as the *Dynamic Resource Constrained Reward Collection* (DRCRC) problem. There are L resources and the decision-maker is initially endowed with initial capacities $C \in \mathbb{R}_+^L$ for the resources.

In each period t, an opportunity arises and each opportunity is characterized by a class $\theta \in \Theta$, where Θ is a finite set of types.[2] Classes are drawn independently from a distribution $p \in \Delta(\Theta)$. Upon observing an opportunity, the decision-maker takes an action $a \in \mathcal{A}$, where $\mathcal{A}$ is the set of feasible actions. Upon taking an action a, the decision-maker collects a reward that depends on the opportunity class θ, the action a, and an idiosyncratic shock ε. Shocks lie in a space $\mathcal{E}$ and are drawn independently from a distribution $f \in \Delta(\Theta)$.[3] Shocks are revealed to the decision-maker after an action is taken and are meant to capture exogenous factors that are idiosyncratic to the opportunity. We denote by $r : \Theta \times \mathcal{A} \times \mathcal{E} \to \mathbb{R}$, the reward function, where $r(\theta, a, \varepsilon)$ denotes the reward when class is θ, the action is a, and the shock is ε. Taking an action consumes resources and we assume that the amount of resources consumed depends on the opportunity class θ and the value of the shock ε. We denote by $y : \Theta \times \mathcal{A} \times \mathcal{E} \to \mathbb{R}^L$, the vector-valued resource consumption function. In particular $y_l(\theta, a, \varepsilon)$ represents the consumption of resource l if class θ arrived, the decision maker chose an action a, and the shock was ε.

To ensure that the problem is feasible, we assume there is a null action a_0 in $\mathcal{A}$ that consumes no resources and generates no reward. That is, for every opportunity class θ and

[2]Here, we assume finiteness of the set of classes for expository purposes. Our results readily extend to more general spaces.

[3]In the general model, to simplify notation, we make the probability distribution of ε independent of the action and the class but any dependencies may be modeled by modifying reward and consumption functions, and in some applications, it might be convenient to consider some dependence.

idiosyncratic shock ε, we have $r(\theta, a_0, \varepsilon) = 0$ and $y_l(\theta, a_0, \varepsilon) = 0$ for every resource l.

We denote the history up to time $t-1$ as $\mathcal{H}_{t-1} = \{\theta_s, a_s, \varepsilon_s\}_{s=1}^{t-1}$. We let Π denote the set of all non-anticipating policies, i.e., the set of policies such that the action at time t, a_t, depends on the observed class of the opportunity in time t and the history up to (and including) time $t-1$. That is, for a policy π, $a_t = a_t^\pi(\theta_t, \mathcal{H}_{t-1})$[4]. The decision maker's objective is to choose a policy $\pi \in \Pi$ that maximizes her expected rewards earned during the horizon. Taking into account that the consumption's constraints must hold almost surely, the stochastic optimization formulation of the decision-maker may be written as follows.

$$
\begin{aligned}
J^*(C,T) = \quad & \sup_{\pi \in \Pi} \ \mathbb{E}\left(\sum_{t=1}^{T} r(\theta_t, a_t^\pi, \varepsilon_t) \right) \\
& \text{s.t} \quad \sum_{t=1}^{T} y_l(\theta_t, a_t^\pi, \varepsilon_t) \leq C_l, \ \forall l \in [L] \ \ \text{(a.s.)}\,,
\end{aligned}
\tag{3.1}
$$

where $[L]$ denotes the set $\{1, \ldots, L\}$.

Note that in general, this is a dynamic program with potentially a high number of dimensions and the curse of dimensionality precludes solving this problem to optimality. Given this, various heutistics can considered and their performance can be assessed through the resulting optimality gap

$$J^*(C,T) - J^\pi(C,T),$$

where $J^\pi(C,T)$ represents the expected reward obtained by the decision-maker if policy π is implemented. We refer to the expression above as the *reward loss* of the heuristic given by π.

3.5 Certainty Equivalent Heuristic

As it was mentioned in Section 1.4, the optimal solution of the stochastic formulation of the DRCRC is not easy to compute. A common and central heuristic in the theory of dynamic decision-making under uncertainty is based on the *certainty equivalent principle*: replace quantities by their expected values and take the best actions given the current history. Specifically, at each point of time t, we solve an optimization problem obtained by using the history up to $t-1$ and replacing the random quantities in problem (3.1) by their expectations.

That is, if we denote by Φ the set of all class-dependent probability distributions $\phi : \Theta \to \Delta(\mathcal{A})$, and by $\rho \in \mathbb{R}_+^L$ a non-negative parameter representing the vector of available inventory divided by the number of remaining periods, at time t we solve the following

[4] To simplify notation we will write a_t^π to refer to the action taken at time t given the policy π

parametric programming problem for $\rho = \rho_t$, which we refer to as the *deterministic problem*:

$$\begin{aligned} J(\rho) = \max_{\phi\in\Phi} \quad & \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \mathbb{E}_\varepsilon\left(r(\theta, a, \varepsilon)\right) \mathrm{d}\phi_\theta(a) \\ \text{s.t} \quad & \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \mathbb{E}_\varepsilon\left(y_l(\theta, a, \varepsilon)\right) \mathrm{d}\phi_\theta(a) \le \rho_l, \ \forall l \in [L]. \end{aligned} \tag{3.2}$$

Because the distributions of classes and shocks are independent and identically distributed (i.i.d.) and we allow for randomized actions, we can restrict attention without loss to static, randomized controls in the deterministic problem. Fixing the parameter ρ, for each class $\theta \in \Theta$ and a set of actions $A \in \mathcal{A}$, the decision variable $\phi_\theta(A)$ gives the probability of choosing an action $a \in A$ given that the arrival belongs to class θ.

In what follows, we assume that problem (3.2) admits an optimal solution. We denote by ϕ^*_ρ an optimal solution when the parameter is ρ. We will provide sufficient conditions for existence of an optimal solution in Section 3.7.[5]

The Certainty Equivalence Principle leads to a natural heuristic for the decision-maker: at each point in time t, choose actions according to a solution $\phi^*_{\rho_t}$ to (3.2) with $\rho_t = c_t/(T-t+1)$ and c_t the capacity remaining at beginning of time t. We call this heuristic the *Certainty Equivalent Heuristic* (CE) and denote the corresponding policy by π^{CE}. The heuristic is formally presented in Algorithm 3.

The certainty equivalence heuristic has been extensively studied in the literature for specific applications. Our aim is to characterize its performance for the broader class of DRCRC problems. Thus, as the family of DRCRC problems encompasses a large number of applications that have been studied in the literature separately, by analyzing the performance of the CE heuristic, we shall recover some already known results and, in the process, obtain new results for other applications, while highlighting very general sufficient conditions to ensure "good" performance of the CE heuristic.

Remark. An alternative "fluid" heuristic involves solving the deterministic problem once at the beginning of the horizon with $\rho_1 = C/T$ and then implementing the static control ρ_1 until capacity runs out. This static control is more computationally tractable as it does not need to resolve the deterministic problem, but typically yields worse performance (see, Appendix C.1.1 for further details).

[5]When an optimal solution does not exist, it suffices for our purpose to work with a $1/T$ approximately optimal solution.

Algorithm 3 Certainty Equivalence Heuristic (CE)

Initialize $c_1 \leftarrow C$,
for $t = 1$ to T **do**
  $\rho_t \leftarrow c_t/(T - t + 1)$
  $\phi^*_{\rho_t} \leftarrow$ optimal solution of Problem 3.2 with $\rho = \rho_t$.
  observe the opportunity class θ_t
  draw an action a_t with probability $\phi^*_{\rho_t}(\theta_t, a_t)$
  if $y(\theta_t, a_t, \varepsilon) \leq c_t,\ \forall \varepsilon \in \mathcal{E}$, **then**
    choose the action a_t
    observe the shock ε_t
    $c_{t+1} \leftarrow c_t - y(\theta_t, a_t, \varepsilon_t)$
  else
    choose the null action a_0
    $c_{t+1} \leftarrow c_t$

3.6 Bound on the cumulative reward loss for DRCRC problems

In this section we study the performance of the CE heuristic for the general class of DRCRC problems. To this end, we first introduce some definitions and conditions on the primitives.

We will assume that the reward and consumption functions are bounded. For a vector $x \in \mathbb{R}^n$, we denote by $\|x\| = \left(\sum_{\mathrm{i}=1}^n x_\mathrm{i}\right)^{1/2}$ its ℓ^2-norm and denoted by $\|x\|_\infty = \max_\mathrm{i} |x_\mathrm{i}|$ its ℓ^∞-norm.

Assumption 1. *The following hold:*

1. *There exists $\bar{r}_\infty \in \mathbb{R}_{++}$ such that $|r(\theta, a, \varepsilon)| \leq \bar{r}_\infty$ for all $\theta \in \Theta, a \in \mathcal{A}$, and $\varepsilon \in \mathcal{E}$.*
2. *There exist $\bar{y}_2, \bar{y}_\infty \in \mathbb{R}_{++}$ such that $\|y(\theta, a, \varepsilon)\| \leq \bar{y}_2$ and $\|y(\theta, a, \varepsilon)\|_\infty \leq \bar{y}_\infty$ for all $\theta \in \Theta, a \in \mathcal{A}$, and $\varepsilon \in \mathcal{E}$.*

Recall that ρ_1 is the vector of initial inventory divided by the amount of periods to consider. We assume that, in the neighborhood of ρ_1, the optimal objective value of the deterministic problem $J(\rho)$ is locally smooth as well as that the consumption constraints are binding.

Assumption 2. *There exist $\delta, K \in \mathbb{R}_{++}$ with $\delta < \min_{j \in [L]} (\rho_1)_j$ such that for every ρ with $\|\rho - \rho_1\| < \delta$, then it holds that:*

1. *The function $J(\rho)$ satisfies $J(\rho) \geq J(\rho_1) + \nabla J(\rho_1)(\rho - \rho_1) - \frac{K}{2}\|\rho - \rho_1\|^2$.*
2. *The optimal solution ϕ^*_ρ satisfies $\mathbb{E}_{\theta \sim p, a \sim \phi^*_\rho, \varepsilon \sim f}(y(\theta, a, \varepsilon)) = \rho$.*

We will refer to inequality given in Assumption 2.1 as $J(\rho)$ admitting a *K-lower downward quadratic* (K-LDQ) *envelope* in $\mathcal{N}(\rho_1, \delta)$, where we denote by $\mathcal{N}(\rho_1, \delta)$ the ball of radius δ centered at ρ_1[6]. See Figure 3.1 (a) for an example of a function admitting a K-LDQ envelope and its envelope. This condition is a weaker and local notion of the K-strongly smooth condition for concave functions, which requires the inequality in Assumption 2.1 to hold for every pair of parameters ρ, ρ'. A sufficient condition for $J(\rho)$ to admit a K-LDQ envelope is that its gradient is *locally K-Lipschitz continuous* at ρ_1 for all $\rho \in \mathcal{N}(\rho_1, \delta)$, that is,

$$\|\nabla J(\rho_1) - \nabla J(\rho)\| \leq K \|\rho_1 - \rho\| \,.$$

See Lemma 3.4 in [33] for a proof of the previous fact.

We are now ready to state our performance bound of the CE heuristic under Assumptions 1 and 2. Specifically, in Theorem 6 we state that the reward loss of the heuristic given by CE is on the order of $\log T$, giving an explicitly expression for the bound.

Theorem 6 Let J^{CE} be the expected performance of Algorithm 3. Then, under Assumptions 1 and 2, the reward loss satisfies

$$J^* - J^{\mathrm{CE}} \leq \frac{1}{2} \bar{y}_2^2 K \log T + \left[\Psi + 14 \frac{\bar{y}_\infty^2}{\delta^2} \right] J(\rho_1),$$

where $\Psi = \frac{\bar{y}_\infty}{\bar{\rho}_1 - \delta}$ and $\bar{\rho}_1$ the smallest component of vector ρ_1.

We note that the result above applies across all DRCRC problems, and only requires Assumption 2. Consider a regime in which C and T are scaled proportionally, i.e., $C = \rho_1 T$ for some $\rho_1 \in \mathbb{R}^L_{++}$. Theorem 6 implies that, in such a regime, the CE heuristic is asymptotically optimal in the sense that $J^{\mathrm{CE}}/J^* \to 1$ as $T \to \infty$ because the reward collected by the CE heuristic grows as $T \to \infty$. Furthermore, the optimality gap is of order $K \log T$ if $K > 0$ and of order 1 if $K = 0$. In other words, we already see appearing clear distinction among DRCRC problems driven by the value of K in Assumption 2.

At a more detailed level, the dependency on the number of resources L enters our bound indirectly via the constants $\bar{y}_2$, $\bar{y}_\infty$, and K. Interestingly, when resource consumption is uniformly bounded, i.e., $\bar{y}_\infty < \infty$, the dependency on the number of resources is mostly driven by $\bar{y}_2^2$. While in the worst case we could have $\bar{y}_2^2 = \Omega(L)$, in many settings of interest, one will have $\bar{y}_2^2 = O(1)$ and obtain bounds that are independent of the number of resources. This could happen, for example, if every opportunity consumes only a finite subset of resources.

The proof of the theorem can be found in Appendix C.2. The main idea of the proof of Theorem 6 is to analyze the performance of the CE heuristic up to the stopping time τ,

[6] More specifically, $\mathcal{N}(x, r)$ represents the set of all points at distance less or equal to r from x, where the distance is the one induced by the ℓ^2-norm: $\mathcal{N}(x, r) = \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$.

where τ is the first time that a resource is close to depletion or the ratio of capacity to time remaining ρ_t leaves the ball $\mathcal{N}(\rho_1, \delta)$ defined in Assumption 2. Using that the deterministic problem gives an upper bound on the optimal value of the stochastic problem, that is, $J^* \le TJ(C/T)$ (see, e.g., [66] for specialized argument), we can bound the reward loss as follows

$$J^* - J^{\mathrm{CE}} \le TJ(\rho_1) - J^{\mathrm{CE}} \le \mathbb{E}\left(\sum_{t=1}^{\tau} J(\rho_1) - \sum_{t=1}^{\tau} r(\theta_t, a_t^{\pi^{\mathrm{CE}}}, \varepsilon_t)\right) - \mathbb{E}\left(\sum_{t=\tau+1}^{T} J(\rho_1)\right), \quad (3.3)$$

where $a_t^{\pi^{\mathrm{CE}}}$ denotes the action taken by the CE heuristic and the second inequality follows because $r(\theta_t, a_t^{\pi^{\mathrm{CE}}}, \varepsilon_t) \ge 0$ since the null action a_0 is feasible. The second term of the right-hand side can be written as $\mathbb{E}(T-\tau)J(C/T)$, which is on the order of $O(1)$, as we establish in Lemma 16. This follows because, under the CE heuristic, the ratio ρ_t behaves like a martingale by Assumption 2.2 and, as result, the heuristic never runs out of resources nor ρ_t leaves the ball $\mathcal{N}(\rho_1, \delta)$ too early. The first term is shown to be of order $O(\log T)$. To see this, note that up to time τ actions are not constrained by resources and the expected reward at period t satisfies $\mathbb{E}\left(r(\theta_t, a_t^{\pi^{\mathrm{CE}}}, \varepsilon_t) \mid \rho_t\right) = J(\rho_t)$ because the CE heuristic takes actions according to $\phi^*_{\rho_t}$. Therefore, using Assumption 2.1 we can upper bound the first term by $\nabla J(\rho_1)\mathbb{E}\left(\sum_{t=1}^{\tau}(\rho_1 - \rho_t)\right) + K/2\mathbb{E}\left(\sum_{t=1}^{\tau} \|\rho_1 - \rho_t\|^2\right)$. The first term is zero because ρ_t behaves like a martingale, while the second term can be bounded using that martingale differences are orthogonal. Putting everything together, we then conclude that $J^* - J^{\mathrm{CE}} = O(K \log T)$.

In what follows, we analyze the reward loss depending on whether the set of actions is finite or not. Specifically, we provide sufficient conditions on the problem primitives for Assumption 2 to be satisfied. These conditions shall yield closed-form expressions for the values of δ and K values. Though we have bounded the reward loss for those problems included in DRCRC, we will show that for an important subfamily of problems, the result is valid with $K = 0$ and, therefore, we recover a constant reward loss instead of order $\log T$.

3.7 Dual problem and sufficient conditions for Assumption 2

Before stating sufficient conditions for Assumption 2, we introduce a dual of Problem 3.2 in which we dualize the consumption constraints. To this end, let $\mu \in \mathbb{R}^L_+$ be the vector of Lagrange multipliers associated with the consumption constraints of Problem 3.2. Let $\bar{r} : \Theta \times \mathcal{A} \to \mathbb{R}_+$ denote the expected reward function, i.e., $\bar{r}(\theta, a) = \mathbb{E}_\varepsilon(r(\theta, a, \varepsilon))$. In the same way, for each $l \in [L]$, let $\bar{y} : \Theta \times \mathcal{A} \to \mathbb{R}^L_+$ denote the expected resource consumption function, i.e., $\bar{y}(\theta, a) = \mathbb{E}_{\varepsilon}(y(\theta, a, \varepsilon))$. Then, the Lagrangian function is given by

$$\begin{aligned}\mathcal{L}(\phi, \mu) &= \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \bar{r}(\theta, a)\, \mathrm{d}\phi_\theta(a) + \mu^\top \left(\rho - \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \bar{y}(\theta, a)\, \mathrm{d}\phi_\theta(a)\right) \\ &= \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \left(\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\right) \mathrm{d}\phi_\theta(a) + \mu^\top \rho.\end{aligned}$$

Define, for each $\theta \in \Theta$, the function $g_\theta : \mathbb{R}^L \to \mathbb{R}$

$$g_\theta(\mu) = \sup_{a\in\mathcal{A}} \left\{\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\right\}. \tag{3.4}$$

The Lagrange dual function, for fixed $\rho \geq 0$, is given by

$$\Psi_\rho(\mu) = \sup_{\phi_a\in\Delta(\mathcal{A})} \mathcal{L}(\phi, \mu) = \mu^\top \rho + \sum_{\theta\in\Theta} p_\theta \sup_{a\in\mathcal{A}} \left\{\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\right\} = \mu^\top \rho + \sum_{\theta\in\Theta} p_\theta g_\theta(\mu),$$

where the second equality follows from optimizing point-wise over actions. The dual problem of Problem 3.2 is then given by

$$\inf_{\mu\in\mathbb{R}^L_+} \Psi_\rho(\mu). \tag{3.5}$$

In what follows, we will assume that the dual solution for $\rho = \rho_1$, namely μ^1, is interior; that is, $\mu^1 > 0$. We introduce this assumption to guarantee that all resource constraints are binding for $\rho = \rho_1$, but we believe our results still hold without this assumption.

3.7.1 Continuum of actions

We first assume that the set of $\mathcal{A}$ is a continuum. In this case we will give a closed-form expression for values $K > 0$ and δ for which Assumption 2 holds and we obtain, under certain conditions that we explain below, a reward loss of the order of $O(K \log T)$.

We now present a condition under g_θ that is sufficient to ensure that the deterministic problem has zero duality gap for every positive parameter ρ. Furthermore, we prove that the problem admits an optimal, deterministic solution. The latter implies that when the set of actions is a continuum, under assumption SC 1, randomization in the deterministic problem is not needed.

SC 1. *For each $\theta \in \Theta$, $g_\theta(\mu)$ is differentiable in μ, and for every $\mu \geq 0$, $g_\theta(\mu)$ is achieved for an action $a \in \mathcal{A}$.*

Proposition 7. *Under Assumptions 1 and SC 1, strong duality holds, i.e., $J(\rho) = \inf_{\mu\in\mathbb{R}^L_+} \Psi_\rho(\mu)$ for all $\rho > 0$. Furthermore, $J(\rho)$ admits a deterministic optimal solution for all ρ.*

A proof is provided in Appendix C.3.1. In addition to condition SC 1, we need to make another regularity assumption over the function g_θ, stated below, to ensure Assumption 2 holds for this particular cases of DRCRC.

SC 2. *There exists positive real numbers κ and $\nu \le \underline{\mu}$ with $\underline{\mu} = \min_{l \in [L]} \mu_l^1$ such that for all $\theta \in \Theta$ and $\mu \in \mathcal{N}(\mu^1, \nu)$, $g_\theta(\mu)$ satisfies*

$$g_\theta(\mu) \ge g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top (\mu - \mu^1) + \frac{\kappa}{2} \|\mu - \mu^1\|^2 . \tag{3.6}$$

In what follows we will refer to property (3.6) as g_θ admitting a *κ-lower upward quadratic* (κ-LUQ) *envelope* in $\mathcal{N}(\mu^1, \nu)$. See Figure 3.1 (b) for an example of a function admitting a κ-LUQ and its envelope. Note that admitting a κ-LUQ envelope is a weaker and local notion of the κ-strongly convex condition, which requires (3.6) to hold for every pair of dual variables μ, μ'.

Lemma 7. *Suppose that Assumptions SC 1 and SC 2 hold. If $\mu^1 > 0$, Assumption 2 holds with $K = 1/\kappa$ and $\delta = (\nu\kappa)/2$.*

A proof is provided in Appendix C.3.2. We obtain the following result as a corollary.

Corollary 2. *Suppose that Assumptions 1, SC 1, and SC 2 hold and that $\mu^1 > 0$. Then the reward loss of the certainty equivalent heuristic is of order $O(\kappa^{-1} \log T)$ for DRCRC problems with a continuum set of actions.*

Geometric interpretation. The deterministic function $J(\rho)$ can be easily shown to be concave and non-decreasing. Assumption SC 2 states that the dual function admits a κ-LUQ envelope in a neighborhood of the Lagrange multiplier μ^1. In Figure 3.1 (b) we represent the dual function for a one-resource problem. By duality, this allows us to prove the smoothness condition on the deterministic function stated in the first statement of Assumption 2, which is represented in Figure 3.1 (a) and consists of $J(\rho)$ admitting a K-LDQ envelope.

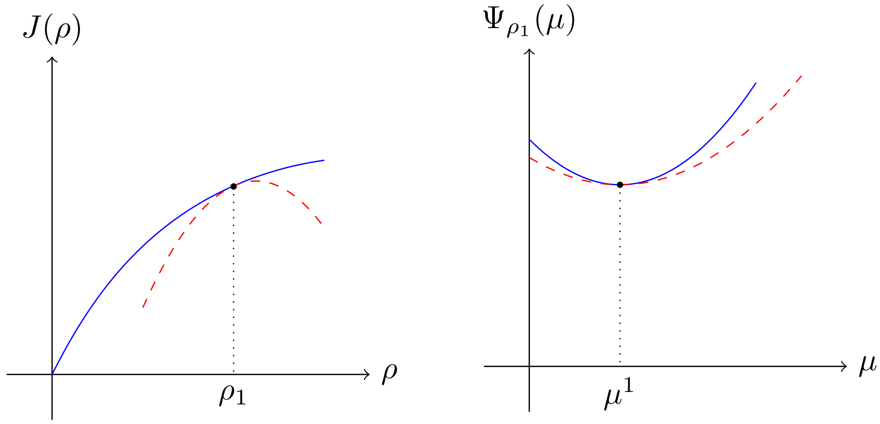
Sufficient conditions on the primitives

The assumptions presented above are stated in terms of $g_\theta(\mu)$, which is a derived object, and, in general, might not be easy to verify. We now present sufficient conditions on the primitives of the problem for Assumptions SC 1 and SC 2 to hold, which, in turn, imply Assumption 2 and allows us to recover a reward loss of order $O(\log T)$ for DRCRC problems with a continuum set of actions.

CA 1. *For each $\theta \in \Theta$, $\bar{r}(\theta, a)$ is upper-semicontinuous in a.*

CA 2. *For each $\theta \in \Theta$, $\bar{y}(\theta, a)$ is continuous in a.*

CA 3. *For each $\mu \ge 0$, the set $A_\mu^* = \left\{a^* \in \mathcal{A} : a^* \in \arg\max_a \{\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\}\right\}$ is not*



(a) Deterministic proxy function in blue. Its 0.9-LDQ envelope in red.

(b) Lagrangian dual function in blue. Its 0.3-LUQ envelope in red.

Figure 3.1: Deterministic proxy J and its Lagrangian dual function, and their envelopes.

empty. Furthermore, the set $\{\bar{y}(\theta, a^*) : a^* \in A^*_\mu\}$ *is a singleton.*

Proposition 8. *If the set of actions $\mathcal{A}$ is compact and conditions CA 1, CA 2 and CA 3 hold, then condition SC 1 is fulfilled.*

The proof of Proposition 8 follows directly from Corollary 4 in [108] and as omitted. It follows from Proposition 8 that if $\mathcal{A}$ is compact and CA 1, CA 2 and CA 3 hold, then there is a solution to the deterministic problem 3.2.

For each $\theta \in \Theta$, let us denote by a^1_θ the feasible action that maximizes $\bar{r}(\theta, a) - (\mu^1)^\top \bar{y}(\theta, a)$. We will assume the following extra conditions in order to bound the reward loss.

CA 4. *For each $\theta \in \Theta$, the feasible action a^1_θ is interior. That is, there exists a positive number φ such that $\mathcal{N}(a^1_\theta, \varphi) \subseteq \mathcal{A}$ for all θ in Θ.*

CA 5. *$\bar{r}(\theta, \cdot)$ admits a κ_r-LDQ envelope in $\mathcal{N}(a^1_\theta, \varphi)$. That is, for all $\theta \in \Theta$,*

$$\bar{r}(\theta, a) \geq \bar{r}(\theta, a^1_\theta) + \nabla \bar{r}(\theta, a^1_\theta)^\top (a - a^1_\theta) - \frac{\kappa_r}{2} \|a - a^1_\theta\|^2 \quad \forall a \in \mathcal{N}(a^1_\theta, \varphi).$$

CA 6. *There exists a positive vector κ_y such that for all $\theta \in \Theta$ and $a \in \mathcal{N}(a^1_\theta, \varphi)$ the consumption function $\bar{y}(\theta, \cdot)$ satisfies*

$$\bar{y}(\theta, a) \leq \bar{y}(\theta, a^1_\theta) + \nabla \bar{y}(\theta, a^1_\theta)(a - a^1_\theta) + \frac{\kappa_y}{2} \|a - a^1_\theta\|^2 , \tag{3.7}$$

where $\nabla \bar{y}(\theta, \cdot)$ represents the Jacobian matrix.

We will refer to property (3.7) as the consumption function $\bar{y}(\theta, \cdot)$ admitting a *κ_y-upper upward quadratic* (κ_y-UUQ) *envelope* in $\mathcal{N}(a^1_\theta, \varphi)$. Note that the κ_y-UUQ envelope condition is a weaker and local notion of the κ_y-strongly smooth condition for concave functions, which requires (3.7) to hold for every pair of action a, a'. Analogously to the relation made with the lower downward quadratic envelope notion, a sufficient condition for Assumption 6 to hold is that the gradient $\nabla \bar{y}_j(\theta, \cdot)$ is locally Lipschitz continuous for every resource $j \in [L]$.

Given a real-valued matrix A, we denote by $\|A\| = \sigma_{\max}(A)$, where $\sigma_{\max}(M)$ represents the largest singular value of matrix M. Recall that given a real-valued matrix A, its singular values are the square roots of the eigenvalues of matrix $A^\top A$. We are now ready to provide sufficient conditions for Assumption SC 2 to hold.

Lemma 8. *Suppose that Assumptions CA 1-CA 6 hold. Then, if $\mathcal{A}$ is compact, SC 2 holds with $\nu = \kappa\varphi/\sigma$ and $\kappa = \kappa_r + (\nu + \|\mu^1\|)\|\kappa_y\|$ where $\sigma = \min_{\theta\in\Theta} \sigma_\theta$, with σ_θ the minimum singular value of $\nabla \bar{y}(\theta, a^1_\theta)$.*

A proof is provided in Appendix C.3.3. We obtain the following result as a corollary

Corollary 3. *Suppose that Assumption 1 and Assumptions CA 1-CA 6 hold. If the set of feasible actions $\mathcal{A}$ is compact and $\mu^1 > 0$, the reward loss of the certainty equivalent heuristic is on the order of $O(\kappa^{-1} \log T)$ for the DRCRC problems with a continuum set of actions.*

3.7.2 Finite set of actions

When the set of actions $\mathcal{A}$ is finite, we will show that, under condition SC 3 below, the reward loss of the certainty equivalent heuristic is on the order of $O(1)$.

SC 3. *The dual problem 3.5 has a unique solution, μ^1, for $\rho = \rho_1$.*

Note that Problem 3.2, in this special case, can be written as follows

$$
\begin{aligned}
J(\rho) = \quad & \max_{\phi_\theta \in \Delta(\mathcal{A})} \quad \sum_{\theta\in\Theta} p_\theta \bar{r}_\theta \phi_\theta \\
& \qquad\quad \text{s.t} \quad \sum_{\theta\in\Theta} p_\theta \bar{y}_\theta \phi_\theta \le \rho\,,
\end{aligned}
\tag{3.8}
$$

where $\bar{r}_\theta = (\bar{r}(\theta, a))_{a\in\mathcal{A}} \in \mathbb{R}_+^{|\mathcal{A}|}$ is the vector of expected rewards for the different actions and $\bar{y}_\theta = (\bar{y}(\theta, a))_{a\in\mathcal{A}} \in \mathbb{R}_+^{L\times|\mathcal{A}|}$ is the matrix of expected resource consumption. Problem (3.8) is a finite-dimensional linear programming problem. Furthermore, the feasible set is non empty and compact, and therefore there exists an optimal solution. In this case, (3.5) can be thought of as a "partial" dual problem in which we dualize the resource constraints but not the simplex constraint $\sum_{a\in\mathcal{A}} \phi_\theta(a) \le 1$. Then, it follows that if the set of actions is finite, the duality gap is zero and $J(\rho) = \inf_{\mu\in\mathbb{R}_+^L} \Psi_\rho(\mu)$.

Note that by Assumption SC 3, there exists non-degenerate optimal solution of (3.8) for $\rho = \rho_1$. Denote such solution by $\phi^*_{\rho_1}$. Because the dual solution is interior, all resource constraints are binding by complementary slackness and this solution satisfies $\sum_{\theta\in\Theta} p_\theta \bar{y}_\theta(\phi^*_{\rho_1})_\theta = \rho_1$.

Considering the standard form of Problem 3.8 (see Appendix C.1.2 for details), let $B \in \mathbb{R}^{(L+|\Theta|)\times(L+|\Theta|)}$ be the corresponding optimal basis matrix and denote by $B^{-1}_{\rho_1}$ the submatrix of B^{-1} associated to the resource constraints. In the following lemma, we show that Assumption 2 holds under Assumption SC 3 whenever $\mu^1 > 0$.

Lemma 9. *Suppose that Assumption SC 3 holds. Then, if $\mu^1 > 0$, Assumption 2 holds with $K = 0$ and $\delta = \phi^*_{\min}/\|B^{-1}_{\rho_1}\|$, where $\phi^*_{\min} = \min_{\theta\in\Theta, a\in\mathcal{A}}\{\phi^*_\theta(a) : \phi^*_\theta(a) > 0\}$.*

The proof is provided in Appendix C.3.4. A direct result from Lemma 9 is that the reward loss if the action space is finite is on the order of $O(1)$.

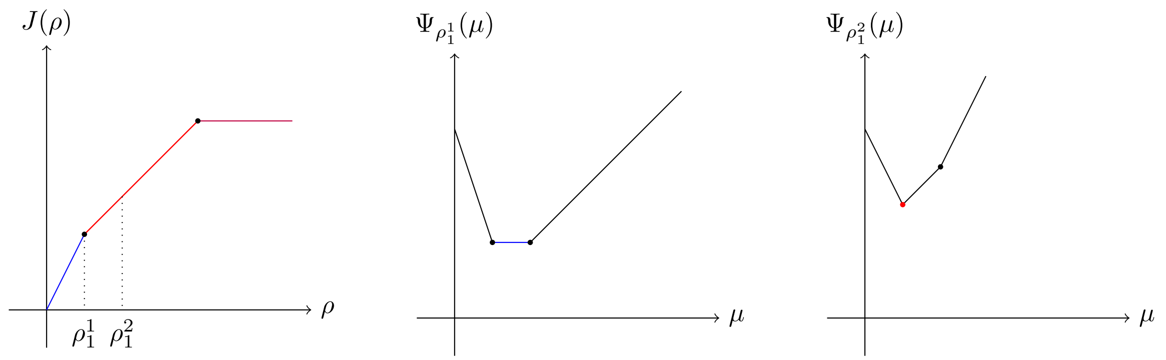
Corollary 4. *Suppose that Assumptions 1 and SC 3 hold and that the dual solution is interior. Then, the reward loss of the certainty equivalent heuristic is bounded by a constant for DRCRC problems with finite actions.*

Geometric interpretation. Let us now provide a geometric interpretation for assumption SC 3 as well as for the deterministic and dual functions for the case of finite actions.

First, note that problem (3.8) is a LP and therefore the deterministic function $J(\rho)$ is a concave piece-wise linear function (see [24] for more details). Moreover, due to the nature of the problem, it will be non-decreasing. In Figure 3.2 (a), the function $J(\rho)$ is plotted for a problem with one resource and two classes. Every optimal dual variables μ for $\Psi_\rho(\mu)$ gives a super-gradient to $J(\rho)$. Therefore, the slope of each straight-line segment is equal to the Lagrange multiplier associated to the consumption constraint, and the corresponding interval gives the values of the right-hand side range for the consumption constraint ρ for which the same dual variable is optimal.

Figure 3.2 (b) and (c) plot the dual function $\Psi_\rho(\mu)$ as a function of μ for two different possible values of the parameter ρ. In Figure 3.2 (b), we take $\rho = \rho_1^1$, a value where $J(\rho)$ has a kink. In this case, the dual problem has an infinite solutions (blue segment in the figure) and every dual solution is a super-gradient of $J(\rho_1)$. Figure 3.2 (c), we plot the dual function at $\rho = \rho_1^2$, a value belonging to an interval where $J(\rho)$ is smooth. There, the set of super-gradients is a singleton and, as a result, the dual optimal solution is unique (red dot in the figure). Thus, Assumption SC 3 is equivalently asking that the parameters ρ_1 lies in the interior of an interval where the deterministic function $J(\rho)$ is smooth.

Note that here, it is not necessary to assume that the dual function admits a κ-LUQ envelope (condition SC 2). As we can see in Figure 3.2 (b), a lower upward quadratic envelope is obtained for free when the optimal dual solution is unique and the action set is finite, because the dual problem is piece-wise linear.



(a) Function $J(\rho)$

(b) Dual function at ρ_1^1. Multiple optimal dual solutions.

(c) Dual function at ρ_1^2. Unique optimal dual solution.

Figure 3.2: Function J and dual function Ψ_ρ for two different parameters.

In the following two sections, we review a set of special cases of the proposed DRCRC class of problems. We divide them based on whether the set of actions is finite or not. For each problem, we show it can be modelled as a DRCRC problem, then provide problem-specific sufficient conditions for our assumptions to hold, and conclude by establishing the implications of Theorem 6. As we will see, our results allow to recover some existing results in the literature as special cases, but also uncover new results as for other classes of problems studied in the literature.

3.8 Notable applications with continuum of actions

In this section we present some applications studied in the literature that consider a continuous space of actions. Specifically, we will describe the network dynamic pricing and dynamic bidding in repeated auctions problems.

3.8.1 Network Dynamic Pricing Problem

The interest on dynamic pricing problems has grown during the last few decades. The design of near-optimal pricing policies that are easy to implement has been studied under several model variants and heuristic polices are widely used in practice by firms. We refer the reader to the review papers and textbook Bitran and Caldentey [27], Talluri and Van Ryzin [129], Gallego et al. [65].

The problem is characterized by a finite time selling horizon. At the beginning of each period, a set of customers arrives, and a decision maker posts prices to maximize his expected revenue. Demands are stochastic and inventories are finite and without replenishment.

Mapping to a DRCRC problem

In this setting, there is a set of N different products to sell during a finite, discrete horizon. An arrival class $\theta \in \Theta$ represents a customer class characterized by their valuation for the products. The set of actions $\mathcal{A} = \mathbb{R}_+^N$ consists of the set of all feasible price vectors to post for the products.

For each customer class θ, the idiosyncratic shock ε represents some private information on the consumer willingness to pay. Then, given θ, posted prices $a \in \mathcal{A} = \mathbb{R}_+^N$ and a realization of the shock ε, we denote by $D(\theta, a, \varepsilon) \in \mathbb{R}_+^N$ the induced vector of demand. The reward function is given by $r(\theta, a, \varepsilon) = a^\top D(\theta, a, \varepsilon)$ and the consumption function is $y(\theta, a, \varepsilon) = A_\theta D(\theta, a, \varepsilon)$. In the latter expression, $A_\theta \in \mathbb{R}^{L \times N}$ is a matrix where $A_{l\theta}^n$ represents the units of resource l needed to serve a customer class θ with a single unit of product n.

Sufficient Conditions for Assumption 2

Let $\bar{D}(\theta, a) = \mathbb{E}_\varepsilon[D(\theta, a, \varepsilon)]$ denote the expected demand and $\bar{r}(\theta, a) = a^\top \bar{D}(\theta, a)$ the corresponding expected reward function. The deterministic problem can be expressed as follows:

$$
\begin{aligned}
J(\rho) = \quad \max_{\phi \in \Phi} \quad & \sum_{\theta \in \Theta} p_\theta \int_{\mathcal{A}} \bar{r}(\theta, a) \mathrm{d}\phi_\theta(a) \\
\text{s.t} \quad & \sum_{\theta \in \Theta} p_\theta \int_{\mathcal{A}} A_\theta \bar{D}(\theta, a) \mathrm{d}\phi_\theta(a) \leq \rho.
\end{aligned}
\tag{3.9}
$$

We map conditions CA 1-CA 6 to sufficient conditions for this particular problem. The following conditions together with $\mu^1 > 0$ and compactness of the set of actions $\mathcal{A}$ are sufficient for Lemma 8 to hold.

- The expected demand function $\bar{D}(\theta, a)$ is continuous in a.
- The expected resource consumption $A_\theta \bar{D}(\theta, a)$ at a maximizer of $(a - A_\theta \mu)^\top \bar{D}(\theta, a)$ is unique.
- For each $\theta \in \Theta$, the price vector maximizing $(a - A_\theta^\top \mu)^\top \bar{D}(\theta, a)$, namely a_θ^1, is interior. That is, there exists a positive number φ such that $\mathcal{N}(a_\theta^1, \varphi) \subseteq \mathcal{A}$ for all θ in Θ.
- The expected revenue function $\bar{r}(\theta, \cdot)$ admits a κ_r-LDQ envelope in $\mathcal{N}(a_\theta^1, \varphi)$.
- There exists a positive vector κ_y such that the expected demand function $\bar{D}(\theta, \cdot)$ admits a κ_y-UUQ envelope in $\mathcal{N}(a_\theta^1, \varphi)$.

In particular, under the conditions above, from Lemma 7, Assumption 2 holds with $K = 1/\kappa$ and $\delta = (\nu\kappa)/2$, where $\kappa = \kappa_r + (\nu + \|\mu^1\|)\|\kappa_y\|$, $\nu = \kappa\varphi/\sigma$, and σ is a lower bound on the minimum singular value of $A_\theta \nabla \bar{D}(\theta, a_\theta^1)$. Therefore, by Corollary 2, the revenue loss of the certainty equivalent heuristic is on the order of $O(\log T)$ for the network dynamic pricing problem with a continuum set of feasible prices. Another implication of our result is that the optimal pricing policy associated with the deterministic proxy is deterministic and the decision maker does not need to randomize over posted prices.

It is worth mentioning that in the literature this problem is typically analyzed in the demand space, i.e., for each class θ, the decision variables are the expected demands $\lambda = \bar{D}(\theta, a)$ instead of the prices a (see, e.g., Jasin [82]). In many cases, this leads to a more tractable problem because constraints become linear and, under additional conditions, the objective becomes concave. The issue, however, is that restrictive assumptions are needed for this reformulation of the problem to go through. For example, it is typically assumed that the reward function is concave in the demand space and the demand function is invertible. Our result yield similar performance guarantees for the CE heuristic and only requires local smoothness properties of the revenue function, which typically leads to weaker assumptions. This is an important departure from previous work, even when specializing the analysis.

Finally, note that, if we consider a finite set of feasible prices, the proxy deterministic problem problems reduces to a linear program and therefore it is enough to assume the dual solution for $\rho = \rho_1$ is unique and interior to obtain a constant revenue loss for the CE heuristic.

Connection with literature

Several papers study heuristics for the network dynamic pricing problem. For example, Kunnumkal and Topaloglu [95] and Erdelyi and Topaloglu [58] consider a dynamic programming formulation. Both papers consider an airline network in which prices affect the probability of the arrival request. In the former, the authors propose a stochastic approximation algorithm for choosing prices dynamically and prove its convergence. In the latter, they develop two methods for making pricing decisions based on a decomposition of the original dynamic program.

Closer to the specialization of our results are Maglaras and Meissner [103] and Jasin [82]. The former established that a CE heuristic for the pricing problem will always yield an asymptotic weak decrease in the revenue loss compared to a static control. Jasin [82] considers a single customer class and present a certainty equivalent heuristic akin to the CE one. Our general result recovers his bound on the logarithmic revenue loss but our sufficient conditions are weaker than his. Interestingly, Jasin [82] also proves that resolving less frequently yield a revenue loss of the same order and proposes another heuristic that involves solving only a single optimization at the beginning of the selling horizon.

3.8.2 Dynamic Bidding in Repeated Auctions

A special case of a DRCRC problem is the problem faced by a bidder participating in a sequence of repeated auctions to buy opportunities. The bidder has a budget constraint that limits his total expenditure over the horizon and aims to maximize his cumulative utility. This model is mainly motivated by internet advertising markets in which advertisers buy opportunities to display advertisements—an event referred to as an impression— via repeated auctions subject to budget constraints.

Mapping to a DRCRC problem

In this setting, the decision maker is an advertiser. The advertiser is present in the market for T periods and one impression is auctioned per period. Upon the arrival of an impression at time t, the advertiser determines a real-valued valuation $\theta_t \in \Theta \subset (0, \Theta_{\max}]$ for the impression, which is distributed according to $p \in \Delta(\Theta)$, and chooses an action $a_t \in \mathcal{A} = [0, \Theta_{\max}]$ representing his bid in the auction. We denote by C the budget of the advertiser. The shock ε captures all exogenous uncertainty in the auction, such as the bids of the competitors and any potential randomization of the auction. For simplicity, we assume that ε is independent of the buyer's valuation θ but our model can be easily be modified to account for correlation. The auction is characterized an allocation rule $q : \mathcal{A} \times \mathcal{E} \to [0, 1]$ together with a payment rule $m : \mathcal{A} \times \mathcal{E} \to \mathbb{R}$, which determine the probability that the impression is allocated to the advertiser and his expected payment as a function of his bid and the exogenous shock, respectively.

The reward earned by the advertiser and his budget consumption, given a, ε and θ can be expressed as $r(\theta, a, \varepsilon) = (\theta q(a, \varepsilon) - m(a, \varepsilon))$ and $y(\theta, a, \varepsilon) = m(a, \varepsilon)$, respectively. Given an action a, we introduce the interim allocation and interim payment variables defined as follows: $\bar{q}(a) = \mathbb{E}_\varepsilon[q(a, \varepsilon)]$, $\bar{m}(a) = \mathbb{E}_\varepsilon[m(a, \varepsilon)]$.

Sufficient Conditions for Assumption 2

For the particular setting described above, the deterministic problem (3.2) is equivalent to the following problem:

$$
\begin{aligned}
J(\rho) = \quad \max_{\phi \in \Phi} \quad & \sum_{\theta \in \Theta} p_\theta \int_{\mathcal{A}} (\theta \bar{q}(a) - \bar{m}(a)) \, \mathrm{d}\phi_\theta(a) \\
\text{s.t} \quad & \sum_{\theta \in \Theta} p_\theta \int_{\mathcal{A}} \bar{m}(a) \, \mathrm{d}\phi_\theta(a) \le \rho \, .
\end{aligned}
\tag{3.10}
$$

For each $\theta \in \Theta$, let $g_\theta(\mu) = \max_{a \in \mathcal{A}} \{\theta \bar{q}(a) - (\mu + 1)\bar{m}(a)\}$. Assumption SC 1 requires that $g_\theta(\mu)$ is differentiable and that $g_\theta(\mu)$ is achieved by an action. Under these conditions, Proposition 7 implies that strong duality holds and the Problem (3.10) admits a deterministic optimal solution. In this application, we can characterize an optimal bidding strategy in terms of an optimal bidding function for the static auction without budget constraints, which we denote by $\beta : \Theta \to \mathcal{A}$. That is, given an advertiser with valuation θ, the optimal bidding strategy for the static auction (ignoring budget constraints) satisfies

$$\beta(\theta) \in \arg\max_{a \in \mathcal{A}} \{\theta \bar{q}(a) - \bar{m}(a))\} .$$

We have the following result. A proof is provided in Appendix C.4.1.

Proposition 9. *Under Assumption SC 1, an optimal solution of (3.10) is to bid $\beta(\theta/(1+\mu^*))$ when the value is θ, where μ^* is the optimal solution of the dual problem of (3.10).*

If in addition Assumption SC 2 holds, from Lemma 7 we obtain that if $\mu^1 > 0$, Assumption 2 holds with $K = 1/\kappa$ and $\delta = (\nu\kappa)/2$.

Below, we study the particular cases of second-price auction and first-price auction. Specifically, we provide sufficient conditions on the primitives of the problem for conditions SC 1 and SC 2 to be satisfied.

Second-price auctions

In a second-price auction, the bidder with the highest bid wins the auction and pays the second-highest bid. In this case, we reduce the definition of ε to a random variable capturing the maximum bid of the competitors and take $\mathcal{E} = \mathbb{R}_+$. Again, we assume ε is distributed according to f, with density function f'. We assume that ties are broken in favor of the decision maker. The allocation and payment functions are given by $q(a, \varepsilon) = 1_{\{a \geq \varepsilon\}}$ and $m(a, \varepsilon) = \varepsilon 1_{\{a \geq \varepsilon\}}$, respectively.

Suppose that the following conditions hold:

- The distribution of the maximum competing bid f is absolutely continuous and strictly increasing.
- The density f' is locally ξ-Lipschitz continuous with respect to a^1_θ in $\mathcal{N}(a^1_\theta, \varphi)$.

If $\mu^1 > 0$, it is possible to show—see Lemma 12 in Appendix C.1.3—that these conditions are sufficient to apply Proposition 8 and Lemma 8 and, in turn, Corollary 3 holds. This leads to a revenue loss of $O(\log T)$.

First-price auctions

In a first-price auction, the winner is the highest bidder but pays his bid. Again, we reduce the definition of ε to a random variable capturing the maximum bid of the competitors. We assume ε is distributed according to f, with density function f'. The allocation and payment functions are given by $q(a, \varepsilon) = 1_{\{a \geq \varepsilon\}}$ and $m(a, \varepsilon) = a 1_{\{a \geq \varepsilon\}}$, respectively.

Suppose the following conditions hold:

- The distribution of the maximum competing bid f absolutely continuous.
- The function $M(a) = a + f(a)/f'(a)$ is strictly increasing.
- The bid a^1_θ maximizing $\theta\bar{q}(a) - (1+\mu^1)\bar{m}(a)$ is interior. That is, there exists a positive number φ such that $\mathcal{N}(a^1_\theta, \varphi) \subset \mathcal{A}$ for all $\theta \in \Theta$.
- The density f' is locally $\xi-$Lipschitz continuous with respect to a^1_θ in $\mathcal{N}(a^1_\theta, \varphi)$.

Moreover, if we have $\mu^1 > 0$, then it can be proved—see Lemma 13 in Appendix C.1.3—that assumptions CA 1-CA 6 hold and, therefore, by Corollary 3 we obtain an $O(\log T)$ revenue loss.

Connection with literature

While the problem of bidding in repeated auctions with budgets has been studied in the past, to the best of our knowledge, this is the first paper that studies the revenue loss of a certainty equivalent heuristic with resolving for the advertiser's decision problem. The assumption that budgets are large relative to the average price paid in an auction is reasonable for internet advertising markets as bidders participate in a large number of auctions. Many papers consider deterministic fluid approximations and static bidding policies.

Abhishek and Hosanagar [1] study this problem, where the goal is to compute optimal bids for multiple keywords in an advertiser's portfolio. To this end, they propose two bidding policies: one ignores the interaction between keywords, and another that incorporate interaction between keywords, and the paper's focus is the deterministic proxy as opposed to the dynamic program (or the relation between the two). Motivated by ad exchanges, Balseiro et al. [20] introduce a fluid mean-field equilibrium (FMFE) notion to study the strategic outcome of advertisers competing in repeated second-price auctions. They characterize the optimal bidding strategy for an advertiser by using the optimal solution of the static problem with budget constraint and prove that FMFE strategies approximate the rational behavior of the advertisers in large markets. Their analysis, however, is restricted to static policies which attain a revenue loss of order $O(T^{1/2})$ in this setting. Gummadi et al. [77] characterize optimal bidding strategies for bidding in general, repeated auctions under a fluid regime and define a related notion of equilibrium to Balseiro et al. [20].

Fernandez-Tapia et al. [62] study the problem of bidding in repeated auctins when the arrival of requests is a Poisson process and characterize the optimal bidding strategy via its Hamilton-Jacobi-Bellman equation. They show that the optimal bidding strategy can be obtained in almost closed-form by using a fluid limit approximation.

3.9 Notable applications with finite set of actions

In what follows we present some known problems with finite actions that are particular cases of the problem studied in Section 3.7.2. Then, Assumption SC 3 together with the condition $\mu^1 > 0$ are sufficient for Assumption 2 to hold. Furthermore, Lemma 9 and Corollary 4 applied to these particular problems, yield a constant bound to the revenue loss for the certainty equivalent heuristic.

3.9.1 Network Revenue Management Problem

A first notable special case of the DRCRC class with finite set of actions is a classical class of problems in the Revenue Management literature: the Network Revenue Management (NRM) problem. It has been extensively studied in the literature (see, e.g., Talluri and Van Ryzin [129], Gallego et al. [65]) and have also been the basis for various industry solutions.

In the NRM problem, the decision-maker is a firm who is trying to dynamically allocate a limited amount of resources over a finite horizon. Resources are sold to heterogeneous consumers who arrive sequentially over time and belong to different classes depending on their consumption of resources and the fixed fare they pay. The distribution of customer classes is stationary. Upon a customer's arrival, the firm has to decide whether to accept or reject the customer's request. If the customer is accepted and there is enough remaining inventory to satisfy its request, she consumes the resources requested and pays the corresponding fare. Otherwise, no revenue is collected and no resource is used. The decision maker's objective is to maximize the expected revenue earned during the selling horizon.

Mapping to a DRCRC problem

In a NRM problem, a customer class can be captured by $\theta \in \Theta$ and is characterized by their usage of resources and a fixed price they pay for the service. We let f_θ denote the fare associated with class θ. The decision-maker's feasible actions has two values, $\mathcal{A} = \{0, 1\}$, where we represent the action "accept" by 1 and "reject" by 0. In this problem the set of idiosyncratic shocks is empty. The reward if the customer belongs to class θ and the decision-maker chooses an action a is $r(\theta, a) = f_\theta a$. If we denote by $A_\theta = (A_{l\theta})_l \in \mathbb{R}^L$ the consumption vector, where $A_{l\theta}$ is the amount of resource l required to serve a class θ customer, the consumption given that the decision-maker chooses an action a and the customer class is θ is given by $y(\theta, a) = A_\theta a$.

The above leads directly to a special case of our general formulation. In particular, Problem (3.8), in this special case, can be written as follows

$$\begin{aligned} J(\rho) = \quad \max_{\phi_\theta(1) \in [0,1]} \quad & \sum_{\theta \in \Theta} p_\theta \ f_\theta \ \phi_\theta(1) \\ \text{s.t} \quad & \sum_{\theta \in \Theta} p_\theta \ A_\theta \ \phi_\theta(1) \leq \rho \,. \end{aligned} \tag{3.11}$$

Because the set of action consists of $\mathcal{A} = \{0, 1\}$, it is enough to consider decision variables $\phi_\theta(1)$ for all $\theta \in \Theta$ because $\phi_\theta(0) = 1 - \phi_\theta(1)$. Then, for the definition of $\phi^*_{\min}$ involved in δ, we now need to take into account both the non-negative variables and those that are strictly smaller than one. That is, $\phi^*_{\min} = \min_{\theta\in\Theta}\{\phi^*_\theta(1) : \phi^*_\theta(1) > 0\} \wedge \min_{\theta\in\Theta}\{1 - \phi^*_\theta(1) : \phi^*_\theta(1) < 1\}$, where $x \wedge y$ denotes the minimum between x and y. Furthermore, if we denote by A the matrix whose θ^{th} column consists of the vector $p_\theta A_\theta$, B_{ρ_1} represents the submatrix of A in which we only consider the columns associated to variables $0 < {\phi_\theta}^*(1) < 1$, and we set $A_{\rho_1} = B_{\rho_1}$. Then, under Assumption SC 3, if $\mu^1 > 0$, Assumption 2 holds with $K = 0$ and $\delta = \phi^*_{\min}/\|A^{-1}_{\rho_1}\|$, where $\phi^*_{\min} = \min_{\theta\in\Theta}\{\phi^*_\theta(1) : \phi^*_\theta(1) > 0\} \wedge \min_{\theta\in\Theta}\{1 - \phi^*_\theta(1) : \phi^*_\theta(1) < 1\}$.

Connection with literature

As mentioned earlier, the NRM problem has a long history. It was originally proposed in D'Sylva [51], Glover et al. [71] and Wang [131]. The question of approximating optimal performance through simple policies has also received significant attention.

Jasin and Kumar [84] studies a NRM problem with one resource and with arrivals following a Poisson process in continuous time. Discretizing the selling horizon, their model fits a special of the above NRM model. In addition, it is worth noting that their assumptions implies SC 3 and $\mu^1 > 0$, and therefore we recover the main result in Jasin and Kumar [84], in a discrete-time setting, as a special case of Theorem 6.

A series of recent papers, which we review below, study variants of the NRM model. While the problems studied in these papers can be mapped out to be special cases of the DRCRC class of problems, the results developed are of a different nature. The objective is to develop heuristic that depart from the plain CE heuristic and that do not require an assumption akin to Assumption 2 to ensure "good" performance.

Discretizing the time horizon, the setting considered in Reiman and Wang [116] would fit to the one described in this section. In particular, they consider that customers' arrival processes satisfying two conditions: on one hand, they assume that the functional central limit theorem holds for a properly centered and scaled sequence of the arrival processes, and on the other, they make a technical assumption that holds, for example, for an independent renewal processes. They propose a heuristic that resolves the deterministic problem once and they obtain that under their policy the revenue loss is $o(\sqrt{T})$. In a related setting, Bumpensanti and Wang [34] assume that arrivals follow a Poisson process and show that the CE heuristic has a $\Theta(\sqrt{T})$ revenue loss in the general case (without the nondegeneracy assumption). They also propose a heuristic that has a $O(1)$ revenue loss. The idea is to only re-solve the deterministic problem a few selected times, using the approach of Reiman and Wang [116] recursively. More recently, in Vera and Banerjee [130] the authors consider the NRM problem for arrival processes satisfying some quadratic tail bound and they propose an algorithm based on thresholds that allows them to obtain a constant upper bound for the expected revenue loss.

An interesting direction is to explore whether the ideas developed in these papers can be leveraged for the more general DRCRC class of problems.

3.9.2 Choice-Based Network Revenue Management Problems

This problem bears many similarities to the network revenue management problem. In this setting, a firm is trying to dynamically allocate a limited amount of products which are sold to heterogeneous consumers who arrive sequentially and belong to different classes characterized by their product preferences. The key difference with the NRM class is that, upon a customer's arrival, the firm makes an offer and depending on the offer and on the customer's preferences, the consumer selects a single product to buy.

Mapping to a DRCRC problem

In this setting we consider a set of N products, each of them consisting of a set of resources. Product n is priced at m_n. Given a customer class θ and a product n, we denote by $A^n_{l\theta}$ the amount of resource l needed to serve product n to customer θ. The action set is given by $\mathcal{A} = 2^{\{1,\dots,N\}}$, where an action $a \in \mathcal{A}$ represents a set of products to be offered to a consumer.

For each action a and customer class θ, we define the shock random vector $\varepsilon_{\theta a} \in \mathcal{E} = \big\{\varepsilon \in \{0,1\}^N : \sum_{n\in[N]} \varepsilon^n \leq 1\big\}$, where its nth component $\varepsilon^n_{\theta a}$ is 1 if and only if the customer selects product n from the offer a. Then, $\varepsilon_{\theta a} \sim \text{Multinomial}(1, g_{\theta a})$ where $g^n_{\theta a}$[7] is the probability of the consumer choosing product n given that his class is θ and the action taken is a.

Given the consumer class θ, the action a, and the shock realization $\varepsilon_{\theta a}$, the reward function is given by $r(\theta, a, \varepsilon) = \sum_{n\in[N]} m_n \varepsilon^n$ and the consumption function by $y(\theta, a, \varepsilon) = \sum_{n\in[N]} A^n_\theta \varepsilon$. Note that, conditional on the shock ε, the resource consumption does not depend on the action. However, the action affects the distribution of ε.

In the present setting, the deterministic problem can be expressed as

$$\begin{aligned} J(\rho) = \quad \max_{\phi\in\Phi} \quad & \sum_{\theta\in\Theta}\sum_{a\in\mathcal{A}} \phi_\theta(a) p_\theta \sum_{n\in[N]} m_n g^n_{\theta a} \\ \text{s.t} \quad & \sum_{\theta\in\Theta}\sum_{a\in\mathcal{A}} \phi_\theta(a) p_\theta \sum_{n\in[N]} A^n_\theta g^n_{\theta a} \leq \rho. \end{aligned} \tag{3.12}$$

We can define the vector of variables and the associated matrix involved in the constraints of problem (3.12) by setting each column to $p_\theta \sum_{n\in[N]} A^n_\theta g^n_{\theta a}$. Then, under assumption SC 3, if the optimal dual solution of (3.12) for $\rho = \rho_1$ is interior, we can apply Lemma 9 and Assumption 2 holds for $K = 0$ and δ defined as in the statement of the lemma. Therefore,

[7] Without loss of generality, we can apply suitable transformations to the reward and consumption functions to obtain a distribution of ε that is independent of the class and the action.

we obtain a constant bound on the revenue loss of the CE heuristic for the choice-based network revenue management problem.

Connection with literature

The class of choice-based network revenue management problems has appeared under two streams: one corresponds to the called choice-based problem, where an offer is a network (each product is a combination of one or more resources), and the other corresponds to a dynamic assortment optimization problem under constraints, where there is a one-to-one mapping between products and resources.

In the first stream, the single-leg case was introduced by Talluri and Van Ryzin [128] who provided an analysis of the optimal control policy under a general discrete choice model of demand.

Regarding a network setting, Gallego et al. [63] was the first to study a choice-based NRM problem. They consider flexible products in a continuous time horizon and with arrivals following independent Poisson processes. A flexible product consists of a set of alternative products serving a customer class. That is, if a flexible product F is offered by the decision-maker and accepted by the consumer, then the decision-maker assigns him one of the products in F. They present a dynamic programming formulation of the problem and proved that it can be approximated by using an appropriate deterministic control problem. Furthermore, they showed that the latter problem can be solved efficiently by a column generation algorithm for a broad class of consumer choice models (that includes both independent demands and the multinomial logit model).

Liu and Van Ryzin [100] considered a choice-based network RM problem in which each consumer belongs to a market segment (customer type) characterized by a set of products (differents for each segment) in which the consumer is interested and the decision-maker has to decide a set of products to offer in each selling period. Liu and Van Ryzin [100] introduced linear programming formulation of the problem and they show that the revenue obtained under this deterministic program converges to the optimal revenue under the exact dynamic formulation. Their work is based in the deterministic formulation introduced by Gallego et al. [63] but without considering flexible products.

Bront et al. [32] consider the same problems as Liu and Van Ryzin [100] but they allow customer classes to overlap. They compute the consumer behaviour by using an MNL choice model and developed a column generation algorithm to solve the deterministic LP for large-size networks. The associated subproblem is NP-complete and, therefore, they propose a heuristic that is shown to work well on a set of computational examples.

In the other stream, Bernstein et al. [22] consider a dynamic assortment problem in continuous time. In the problem they consider, all products have the same price and for each customer class they compute the probability that a customer belonging to that class chooses a product from the offer according to a MNL model. In the particular case with two customer types and two different products, they characterize the optimal policy and show that the optimal dynamic program may withhold products with low remaining inventory for future customers that are more interested in them. They then propose a threshold-based heuristic for the general problem.

Our formulation shows that a certainty equivalent heuristic admits strong performance guarantees under Assumption SC 3 and also assuming $\mu^1 > 0$. Corollary 4 leads, to the best of our knowledge, to the first such result for this subclass of dynamic assortment optimization problems.

Golrezaei et al. [74] also formulate a related dynamic assortment optimization problem. Their formulation is different in that it focuses on arbitrary, possibly adversarial, sequences of customer arrivals.

3.9.3 Stochastic Depletion Problems

This problem, introduced by Chan and Farias [39], is similar to the Choice-based NRM problem but now the consumption matrix is a random variable and the reward function depends on the realized amount of resource consumed. More specifically, after a customer arrives, the seller chooses an offer and then an amount of each resource is consumed. The revenue earned by the seller depends on the resource consumption.

Mapping to a DRCRC problem

Arrival classes corresponds to customer classes, which encode the consumers' preferences. The set of actions is a set of offers, each of them consisting on a collection of products. After a customer class θ arrives the seller chooses an offer a from the actions set $\mathcal{A}$ and then $y_l(\theta, a, \varepsilon)$ units of resource l are consumed. The reward earned for presenting offer a to customer class θ when the shock is ε is $r(\theta, a, \varepsilon) = g(y(\theta, a, \varepsilon))$ where $g : \mathbb{R}_+^L \to \mathbb{R}_+$.

The deterministic problem is given by

$$\begin{aligned} J(\rho) = \quad & \max_{\phi \in \Phi} \quad \sum_{\theta} p_\theta \sum_{a \in \mathcal{A}} \phi_\theta(a) \mathbb{E}_\varepsilon \left(g(y(\theta, a, \varepsilon)) \right) \\ & \text{s.t} \ \sum_{\theta} p_\theta \sum_{a \in \mathcal{A}} \phi_\theta(a) \mathbb{E}_\varepsilon \left(y(\theta, a, \varepsilon) \right) \leq \rho. \end{aligned} \tag{3.13}$$

Under the conditions that the optimal dual solution of problem (3.13) for $\rho = \rho_1$ is unique and interior, by applying the results of Section 3.7.2, we obtain that the revenue loss for the certainty equivalent heuristic is bounded by a constant for the stochastic depletion problem.

Connection with literature

While this problem was studied in Chan and Farias [39], the focus there is different. In their work, they prove that under some conditions, a myopic policy yields a constant factor of the expected revenue obtained by the optimal policy.

Discretizing the selling horizon, Jasin and Kumar [83] formulate a stochastic depletion problem that fits a a special case of a stochastic depletion problem. In particular, our result directly implies their result of an $O(1)$ revenue loss by considering a frequent resolving heuristic.

3.9.4 Online Matching

Another closely related class of problems is that of online matching. This problem is closely related to the NRM class, but now, opportunities correspond to sets of resources and the decision-maker has to choose any option from the set to maximize her expected revenue earned during the horizon.

Mapping to a DRCRC problem

We have a bipartite graph with resources in one side and classes on the other side. An opportunity of class θ arrives with probability p_θ and the decision-maker need to decide which resource to assign. Calling $\mathcal{L}$ to the set of resources, each class θ has a fare vector $f_\theta \in \mathbb{R}_+^L$ and a resource consumption $A_\theta \in \mathbb{R}_+^L$. The action set is $\mathcal{A} = \mathcal{L} \cup \{0\}$, where the action 0 represents rejecting the request. Given an arrival θ and an action a, the reward is given by $r(\theta, a) = f_{\theta a} 1_{\{a \neq 0\}}$ and the consumption is $y(\theta, a) = A_{\theta a} 1_{\{a \neq 0\}}$. We assume that the bipartite graph is complete. Incomplete graphs can be modelled by setting $f_{\theta j} = -\infty$ if assigning class θ to resource j is not feasible.

In this case, the deterministic problem is given by

$$\begin{aligned} J(\rho) = \quad \max_{\phi \in \Phi} \quad & \sum_{\theta \in \Theta} \sum_{a \in L} \phi_\theta(a) p_\theta f_{\theta a} \\ \text{s.t} \quad & \sum_{\theta \in \Theta} p_\theta \phi_\theta^\top A_\theta \leq \rho, \end{aligned} \tag{3.14}$$

where $\phi_\theta^\top = (\phi_\theta(1), \ldots, \phi_\theta(L))$.

Here, as in the network revenue management problem, we can define the matrix $A \in \mathbb{R}^{L \times L}$ where the θ^{th} column is the vector $p_\theta A_\theta$ and denote A_{ρ_1} the submatrix of A corresponding to the columns associated to the basic variables. Then, if the dual optimal solution of problem (3.14) for $\rho = \rho_1$ is unique and interior, Assumption 2 holds with $K = 0$ and $\delta = \phi^*_{\min}/\|A_{\rho_1}^{-1}\|$, where $\phi^*_{\min} = \min_{\theta \in \Theta, a \in \mathcal{A}} \{\phi^*_\theta(a) : \phi^*_\theta(a) > 0\}$, and the constant bound for the revenue loss is obtained.

Connection with literature

There are several papers on online matching in the literature with the goal to design an algorithm that maximizes the competitive ratio.

The bipartite online matching was introduced by Karp et al. [88] where they consider the case with arrivals in arbitrary order and with the goal of maximizing the total number of matches. To this end, they presented an online algorithm with a competitive ratio of $1-1/e$ and they show that it is the best possible ratio in that setting. This was generalized to vertex-weighted matchings in Aggarwal et al. [5], presenting a $(1-1/e)$-competitive randomized algorithm for general vertex weights. They also show that the same competitive ratio is obtained considering the problem with capacities.

Note that the adwords problem is a particular case of the problem we describe where $A_\theta = f_\theta$. Devanur et al. [49] consider such a problem and provide a reinterpretation of the algorithm on Karp et al. [88] as a randomized primal-dual algorithm. Their analysis also extends to the algorithm presented on Aggarwal et al. [5].

Also the display allocation problem can formulated as an online matching problem, where the initial capacities give the number of impressions each advertiser request and $A_{\theta a} = 1$ because each impression consumes one unit. Feldman et al. [60] studied an stochastic version of this problem assuming that arrivals are i.i.d samples from a known distribution and that the expected number of impressions is integer. They present an algorithm with a approximation factor higher than $1-1/e$. This result was improved by Manshadi et al. [104], who also consider a stochastic setting but they provide a bound for the competitive ratio that holds without having integrality of expectations.

A constant bound for the reward loss was recently obtained by Vera and Banerjee [130] for a different heuristic.

3.9.5 Order Fulfillment Problem

In this section we study a class of problem faced by a retailer who needs to fulfill the orders they receive from different facilities. Specifically, in this problem an order arrives sequentially and a decision-maker has to construct a fulfillment policy to decide from which facility each of the items in the arriving order should be fulfilled.

Mapping to a DRCRC problem

We consider the setting where there are L different items (resources in the general formulation) that could be served from K different facilities. Each facility k is endowed with an inventory $C_k \in \mathbb{R}^L_+$, with C_{kl} representing the initial capacity of item l in facility k, and we consider that facility K is fictitious with infinite initial capacity of all items.

Arrival θ occurs with probability p_θ and corresponds to a request belonging to $\Theta = 2^{\{1,\dots,L\}}$. We assume one order includes at most one unit of each item. Then, $\theta_l = 1$ if and only if item l is included on the offer θ and 0, otherwise.

The decision-maker has to construct a fulfillment policy to decide from which facility $k \in K$ each of the items in θ should be fulfilled in order to maximize his expected revenue. That is, the action set is given by $\mathcal{A} = \{1, \dots, K\}^L$, where given $l \in L$, $a_l = k$ means that item l is served from facility k. Furthermore, serving item l from facility k has an associated fare denoted by f_{lk}.

Given that the order is θ and the decision-maker chooses action a, the consumption of item l in the facility k is $y_{lk}(\theta, a) = 1_{\{a_l = k\}}$, and the reward is given by $r(\theta, a) = \sum_{l\in\theta, k\in K} f_{lk} 1_{\{a_l=k\}}$.

For the particular setting described above, the deterministic problem can be expressed as follows

$$\begin{aligned} J(\rho) = \quad & \max_{\phi\in\Phi} \quad \sum_{\theta\in\Theta}\sum_{a\in\mathcal{A}} p_\theta \phi_\theta(a) \sum_{l\in\theta}\sum_{k\in K} f_{lk} 1_{\{a_l=k\}} \\ & \text{s.t} \quad \sum_{\theta\in\Theta}\sum_{a\in\mathcal{A}} p_\theta \phi_\theta(a) 1_{\{a_l=k\}} \le \rho_{lk}. \end{aligned} \tag{3.15}$$

In this case, as in the previous ones, we can write the constraints of problem (3.15) in matrix form and, thus, obtain an expression for δ involved in the assumption. We do not give an explicit formula for δ to avoid introducing more notation. Moreover, if the dual optimal solution of problem (3.15) for $\rho = \rho_1$ is unique and interior, Assumption 2 holds with $K = 0$ and we recover a constant revenue loss bound for the order fulfillment problem.

Connection with literature

Many different variants of this DRCRC problem have been studied in the literature. For example, papers have considered different objectives to optimize, whether if the model requires a demand forecast of not, multi or single-item approach, among others. We refer the reader Acimovic and Farias [3] for an overview of order fulfillment problems.

Acimovic and Graves [4] modeled the problem as a dynamic program that minimizes the shipping cost plus the expected future costs and they propose an algorithm to approximate it based on the dual variables of the offline deterministic linear programming approximation. They demonstrate that the heuristic works well in practice.

Another paper considering a fulfillment problem is Andrews et al. [9]. The main difference with the model proposed by Acimovic and Graves [4] is that the former allow demand to be adversarial. They propose a primal-dual based algorithm to approximate the dynamic programming value function. They give a bound for the competitive ratio achieved by the algorithm and establish that no online algorithm can achieve a better competitive ratio.

Jasin and Sinha [85] considered the multi-item order fulfillment problem and study the performance of two heuristics derived from the solution of the deterministic linear program where the objective is to minimize the total shipping cost. They investigated the competitive ratio of both heuristics and provide numerical examples showing that one of the algorithms performs very close to optimal.

Some works in the existing literature consider extra constraints related, for instance, to the set of feasible facilities (or resources) from which is it possible to serve an order. Asadpour et al. [12] consider an online allocation problem with equal numbers of types of resources and types of requests with the restriction that a request of type i can be served only by resources of type i and type i+1. If both resources have zero inventory left, then the sale is lost. Their objective is to provide an upper bound on the difference between the performance with and without the above described restriction on fulfillment.

It is worth mentioning that some works consider the order fulfillment problem jointly with the pricing problem (see, e.g., [80], [97]) or jointly with both pricing and display problems (see, e.g., [98]).

3.9.6 Other applications

Although we have mentioned several problems that are special cases of the DRCRC class, there are some other problems in the literature that also belongs to DRCRC class.

For instance, the dynamic knapsack problem and the multisecretary problem, which are both closely related to the NRM problem exposed, are special cases of DRCRC problems. Arlotto and Xie [11] consider the dynamic knapsack problem, which is closely related to the NRM problem exposed. In this case, arrivals correspond to items to allocate, the knapsack can be seen as the resource, and the resource consumption is given by an item's weight. Arlotto and Gurvich [10] consider the multisecretary problem. This problem is a particular case of the our NRM problem in which the arrivals are the candidates, there is one resource with initial capacity equal to the number of secretaries needed, and the reward associated to a candidate is his ability.

It is worth mentioning that it is also possible to consider a combination some of the particular classes of DRCRC problems mentioned (see, e.g., [80],[97],[98]).

Conclusion

Throughout the chapters of this thesis, we address different problems arising mainly from revenue management. However, the problems considered are also of interest in the areas of computer science and economics, and we then make contributions to all these research communities.

Firstly, we introduced the notion of value of observability for a dynamic pricing problem, which is defined as the ratio between the expected revenue of the seller in the settings where she is able to observes the arrival of the customer before offering the price curve and the expected revenue of the seller when she has not have this power. The main contribution is that we provide a constant upper bound for this value, that is independent of all parameters of the problem. This result is robust and surprising, and the main difficulty of the problem arises from the difficulty of solving the unobservable case. In fact, to obtain this bound we use the optimal solution of the observable case and we construct a feasible pricing policy for the unobservable case that allowed us to extract a constant portion of the expected revenue in the observable case. This leads us to think that, although the bound we found is an interesting contribution, a question for future work is whether this bound could be improved. We have also studied lower bounds for the value of observability, but this problem is difficult due to the impossibility of solving, and even modeling, the unobservable case in general. Thus, using dynamic programming for a particular instance, we obtained a lower bound but we believe that another challenge for future work is to be able to improve it. Since the notion of the value of observability is novel, there are also open questions related to whether if the value of observability could be also uniformly bounded for more general settings (e.g., considering more than one buyer), as well as relating this notion to others already studied in the literature, such as the price of discrimination.

Secondly, we proved that the problem of designing posted price mechanisms is equivalent to that of finding stopping rules for optimal stopping problems. More specifically, we showed that if we have a sequential posted price mechanism with a certain approximation guarantee, it can be turned out into a prophet inequality with the same approximation guarantee. To this end, we proved a key technical result of auction theory- that it may be useful also in settings beyond this thesis- that states that for any distribution there is another whose virtual value distributes according to the former distribution. As a corollary of our result, it is possible to translate all known lower bounds from PPMs into prophet inequalities, and back and therefore our contribution is for both auction and optimal stopping theory.

Finally, we introduced a novel general formulation of dynamic optimization problems, that we call dynamic resource constraint reward collection. In particular, we have seen how this class of problems comprises several problems studied separately in the literature. Furthermore, we considered a certainty equivalent heuristic and we studied its performance for the class DRCRC. In particular, this implies that the performance of the heuristic can be proved in a single result for all the known problems that are special cases of a DRCRC problem, recovering the known performance for some problems and providing new results for others. Although it is important how good the heuristic is, it is also important to study the sufficient conditions needed to ensure that good performance. In this way, the thesis introduced an important contribution, stating that some local smooth conditions on a dual function are sufficient to recover the bounds. We also analyzed, not only the geometric interpretation of the conditions, but also some sufficient conditions on the primitives of the model for each of the particular problem exposed. This work opens up the possibility of further generalizations of arguments developed for special cases of DRCRC problems.

Bibliography

[1] Vibhanshu Abhishek and Kartik Hosanagar. Optimal bidding in multi-item multislot sponsored search auctions. *Operations Research*, 61(4):855–873, 2013.

[2] Melika Abolhassani, Soheil Ehsani, Hossein Esfandiari, MohammadTaghi Hajiaghayi, Robert Kleinberg, and Brendan Lucier. Beating 1-1/e for ordered prophets. In *Proceedings of 49th Annual ACM SIGACT Symposium on the Theory of Computing*, STOC '17. ACM, 2017.

[3] Jason Acimovic and Vivek F Farias. The fulfillment-optimization problem. In *Operations Research & Management Science in the Age of Analytics*, pages 218–237. INFORMS, 2019.

[4] Jason Acimovic and Stephen C Graves. Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management*, 17(1):34–51, 2015.

[5] Gagan Aggarwal, Gagan Goel, Chinmay Karande, and Aranyak Mehta. Online vertex-weighted bipartite matching and single-bid budgeted allocations. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1253–1264. SIAM, 2011.

[6] Hyun-soo Ahn, Mehmet Gümüş, and Philip Kaminsky. Pricing and manufacturing decisions when demand is a function of prices in multiple periods. *Operations Research*, 55(6):1039–1057, 2007.

[7] Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM Journal on Computing*, 43(2):930–972, 2014.

[8] Saeed Alaei, Jason Hartline, Rad Niazadeh, Emmanouil Pountourakis, and Yang Yuan. Optimal auctions vs. anonymous pricing. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, FOCS '15, pages 1446–1463. IEEE Computer Society, 2015.

[9] John M Andrews, Vivek F Farias, Aryan I Khojandi, and Chad M Yan. Primal–dual algorithms for order fulfillment at urban outfitters, inc. *Interfaces*, 49(5):355–370, 2019.

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

[11] Alessandro Arlotto and Xinchang Xie. Logarithmic regret in the dynamic and stochastic knapsack problem. *arXiv preprint arXiv:1809.02016*, 2018.

[12] Arash Asadpour, Xuan Wang, and Jiawei Zhang. Online resource allocation with limited flexibility. *Management Science*, 2019.

[13] Yossi Aviv, Yuri Levin, and Mikhail Nediak. Counteracting strategic consumer behavior in dynamic pricing systems. *Consumer-Driven Demand and Operations Management Models*, 131:323–352, 2009.

[14] Yossi Aviv and Amit Pazgal. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management*, 10(3):339–359, 2008.

[15] Yossi Aviv and Gustavo Vulcano. Dynamic list pricing. 2012.

[16] Pablo D. Azar, Robert Kleinberg, and S. Matthew Weinberg. Prophet inequalities with limited information. In *Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 1358–1377. Society for Industrial and Applied Mathematics, 2014.

[17] Moshe Babaioff, Liad Blumrosen, Shaddin Dughmi, and Yaron Singer. Posting prices with unknown distributions. *ACM Trans. Econ. Comput.*, 5(2):13:1–13:20, March 2017.

[18] Moshe Babaioff, Nicole Immorlica, and Robert Kleinberg. Matroids, secretary problems, and online mechanisms. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '07, pages 434–443. Society for Industrial and Applied Mathematics, 2007.

[19] Santiago Balseiro, Omar Besbes, and Dana Pizarro. Dynamic resource constrained reward collection. *Working paper*, 2020.

[20] Santiago R Balseiro, Omar Besbes, and Gabriel Y Weintraub. Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science*, 61(4):864–884, 2015.

[21] Dirk Bergemann and Philipp Strack. Progressive participation. 2019.

[22] Fernando Bernstein, A Gürhan Kök, and Lei Xie. Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, 17(4):538–553, 2015.

[23] Dimitri P Bertsekas. Nonlinear programming. *Journal of the Operational Research Society*, 48(3):334–334, 1997.

[24] Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

[25] David Besanko and Wayne L Winston. Optimal price skimming by a monopolist facing rational consumers. *Management Science*, 36(5):555–567, 1990.

[26] Omar Besbes and Ilan Lobel. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science*, 61(1):92–110, 2015.

[27] Gabriel Bitran and René Caldentey. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–229, 2003.

[28] Liad Blumrosen and Thomas Holenstein. Posted prices vs. negotiations: An asymptotic analysis. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, EC '08, pages 49–49, New York, NY, USA, 2008. ACM.

[29] Simon Board and Andrzej Skrzypacz. Revenue management with forward-looking buyers. *Journal of Political Economy*, 124(4):1046–1087, 2016.

[30] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2009.

[31] Luis Briceño-Arias, José R Correa, and Andres Perlroth. Optimal continuous pricing with strategic consumers. *Management Science*, 63(8):2741–2755, 2017.

[32] Juan José Miranda Bront, Isabel Méndez-Díaz, and Gustavo Vulcano. A column generation algorithm for choice-based network revenue management. *Operations research*, 57(3):769–784, 2009.

[33] Sébastien Bubeck. Convex optimization: Algorithms and complexity. *arXiv preprint arXiv:1405.4980*, 2014.

[34] Pornpawee Bumpensanti and He Wang. A re-solving heuristic with uniformly bounded loss for network revenue management. *arXiv preprint arXiv:1802.06192*, 2018.

[35] Gérard P Cachon and Robert Swinney. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science*, 55(3):497–511, 2009.

[36] Yang Cai and Constantinos Daskalakis. Learning multi-item auctions with (or without) samples. In *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science*, FOCS '17, pages 516–527. IEEE Computer Society, 2017.

[37] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 170–183. ACM, 2017.

[38] René Caldentey, Ying Liu, and Ilan Lobel. Intertemporal pricing under minimax regret. *Operations Research*, 65(1):104–129, 2017.

[39] Carri W Chan and Vivek F Farias. Stochastic depletion problems: Effective myopic policies for a class of dynamic optimization problems. *Mathematics of Operations Research*, 34(2):333–350, 2009.

[40] Shuchi Chawla, Jason D. Hartline, and Robert Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, EC '07, pages 243–251. ACM, 2007.

[41] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the Forty-second ACM Symposium on Theory of Computing*, STOC '10, pages 311–320. ACM, 2010.

[42] Ronald H Coase. Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149, 1972.

[43] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 243–252. ACM, 2014.

[44] José Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Posted price mechanisms for a random stream of customers. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 169–186. ACM, 2017.

[45] José Correa, Patricio Foncea, Dana Pizarro, and Victor Verdugo. From pricing to prophets, and back! *Operations Research Letters*, 47(1):25–29, 2019.

[46] José Correa, Ricardo Montoya, and Charles Thraves. Contingent preannounced pricing policies with strategic consumers. *Operations Research*, 64(1):251–272, 2016.

[47] José Correa, Dana Pizarro, and Gustavo Vulcano. Value of observability in dynamic pricing. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20. ACM, 2020.

[48] Sriram Dasu and Chunyang Tong. Dynamic pricing when consumers are strategic: Analysis of posted and contingent pricing schemes. *European Journal of Operational Research*, 204(3):662–671, 2010.

[49] Nikhil R Devanur, Kamal Jain, and Robert D Kleinberg. Randomized primal-dual analysis of ranking for online bipartite matching. In *Proceedings of the twenty-fourth annual ACM-SIAM symposium on Discrete algorithms*, pages 101–107. SIAM, 2013.

[50] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91:318–333, 2015.

[51] Ewalds D'Sylva. O and d seat assignment to maximize expected revenue. *Unpublished internal report, Boeing Commercial Airplane Company, Seattle, WA*, 1982.

[52] Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. In *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science*, FOCS '17, pages 540–551. IEEE Computer Society, 2017.

[53] Paul Dütting, Felix Fischer, and Max Klimm. Revenue Gaps for Discriminatory and Anonymous Sequential Posted Pricing. *ArXiv e-prints*, July 2016.

[54] Paul Dütting and Robert Kleinberg. Polymatroid prophet inequalities. In *Algorithms-ESA 2015*, pages 437–449. Springer, 2015.

[55] Liran Einav, Chiara Farronato, Jonathan Levin, and Neel Sundaresan. Auctions versus posted prices in online markets. *Journal of Political Economy*, Forthcoming.

[56] Adam N Elmachtoub, Vishal Gupta, and Michael Hamilton. The value of personalized pricing. *Available at SSRN 3127719*, 2018.

[57] Wedad Elmaghraby, Altan Gülcü, and Pınar Keskinocak. Designing optimal preannounced markdowns in the presence of rational customers with multiunit demands. *Manufacturing & Service Operations Management*, 10(1):126–148, 2008.

[58] Alexander Erdelyi and Huseyin Topaloglu. Using decomposition methods to solve pricing problems in network revenue management. *Journal of Revenue and Pricing Management*, 10(4):325–343, 2011.

[59] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. Prophet secretary. *SIAM Journal on Discrete Mathematics*, 31(3):1685–1701, 2017.

[60] Jon Feldman, Aranyak Mehta, Vahab Mirrokni, and Shan Muthukrishnan. Online stochastic matching: Beating 1-1/e. In *2009 50th Annual IEEE Symposium on Foun-*

dations of Computer Science, pages 117–126. IEEE, 2009.

[61] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *Proceedings of the Twenty-sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, pages 123–135. Society for Industrial and Applied Mathematics, 2015.

[62] Joaquin Fernandez-Tapia, Olivier Guéant, and Jean-Michel Lasry. Optimal real-time bidding strategies. *Applied Mathematics Research eXpress*, 2017(1):142–183, 2017.

[63] Guillermo Gallego, Garud Iyengar, Robert Phillips, and Abhay Dubey. Managing flexible products on a network. *Available at SSRN 3567371*, 2004.

[64] Guillermo Gallego, Robert Phillips, and Özge Şahin. Strategic management of distressed inventory. *Production and Operations Management*, 17(4):402–415, 2008.

[65] Guillermo Gallego, Huseyin Topaloglu, et al. *Revenue management and pricing analytics*, volume 209. Springer, 2019.

[66] Guillermo Gallego and Garrett Van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations research*, 45(1):24–41, 1997.

[67] Jérémie Gallien and Shobhit Gupta. Temporary and permanent buyout prices in online auctions. *Management Science*, 53(5):814–833, 2007.

[68] Daniel F Garrett. Intertemporal price discrimination: Dynamic arrivals and changing values. *The American Economic Review*, 106(11):3275–3299, 2016.

[69] Alex Gershkov, Benny Moldovanu, and Philipp Strack. Revenue-maximizing mechanisms with strategic customers and unknown, markovian demand. *Management Science*, 64(5):2031–2046, 2018.

[70] John P. Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. *Journal of the American Statistical Association*, 61:35–76, 1966.

[71] Fred Glover, Randy Glover, Joe Lorenzo, and Claude McMillan. The passenger-mix problem in the scheduled airlines. *Interfaces*, 12(3):73–80, 1982.

[72] Oliver Göbel, Martin Hoefer, Thomas Kesselheim, Thomas Schleiden, and Berthold Vöcking. Online independent set beyond the worst-case: Secretaries, prophets, and periods. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 508–519, 2014.

[73] Kira Goldner and Anna R. Karlin. A prior-independent revenue-maximizing auction for multiple additive bidders. In *Proceedings of the 12th International Conference on Web and Internet Economics - Volume 10123*, WINE 2016, pages 160–173. Springer-Verlag New York, Inc., 2016.

[74] Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. Real-time optimization of personalized assortments. *Management Science*, 60(6):1532–1551, 2014.

[75] Jochen Gönsch, Robert Klein, Michael Neugebauer, and Claudius Steinhardt. Dynamic pricing with strategic customers. *Journal of Business Economics*, 83(5):505–549, 2013.

[76] Joseph F Grcar. A matrix lower bound. *Linear algebra and its applications*, 433(1):203–220, 2010.

[77] Ramakrishna Gummadi, Peter Key, and Alexandre Proutiere. Repeated auctions under budget constraints: Optimal bidding strategies and equilibria. In *the Eighth Ad Auction Workshop*, 2012.

[78] Allan Gut. *Probability: a graduate course*, volume 75. Springer Science & Business Media, 2013.

[79] Mohammad Taghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, volume 7, pages 58–65, 2007.

[80] Pavithra Harsha, Shivaram Subramanian, and Joline Uichanco. Dynamic pricing of omnichannel inventories: Honorable mention—2017 m&som practice-based research competition. *Manufacturing & Service Operations Management*, 21(1):47–65, 2019.

[81] Theodore P. Hill and Robert P. Kertz. Comparisons of stop rule and supremum expectations of i.i.d. random variables. *The Annals of Probability*, 10(2):336–345, 1982.

[82] Stefanus Jasin. Reoptimization and self-adjusting price control for network revenue management. *Operations Research*, 62(5):1168–1178, 2014.

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

[84] Stefanus Jasin and Sunil Kumar. Analysis of deterministic lp-based booking limit and bid price controls for revenue management. *Operations Research*, 61(6):1312–1320, 2013.

[85] Stefanus Jasin and Amitabh Sinha. An lp-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Operations Research*, 63(6):1336–1351, 2015.

[86] Shlomo Kalish. Monopolist pricing with dynamic demand and production cost. *Marketing Science*, 2(2):135–159, 1983.

[87] Thomas Kämpke, Franz Josef Radermacher, et al. Income modeling and balancing. *Lecture Notes in Economics and Mathematical Systems*, 2015.

[88] Richard M Karp, Umesh V Vazirani, and Vijay V Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 352–358, 1990.

[89] Robert P. Kertz. Stop rule and supremum expectations of i.i.d. rando mvariables: A complete comparison by conjugate duality. *Journal of Multivariate Analysis*, 19:88–112, 1986.

[90] Olga Kharif. Supermarkets offer personalized pricing. *Bloomberg*, November 2013.

[91] Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the Forty-fourth ACM Symposium on Theory of Computing*, STOC '12, pages 123–136. ACM, 2012.

[92] Mirko Kremer, Benny Mantin, and Anton Ovchinnikov. Dynamic pricing in the presence of myopic and strategic consumers: Theory and experiment. *Production and Operations Management*, 26(1):116–133, 2017.

[93] Ulrich Krengel and Louis Sucheston. Semiamarts and finite values. *Bull. Amer. Math. Soc.*, 83:745–747, 1977.

[94] Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. *Adv. in Probability*, 4:197–266, 1978.

[95] Sumit Kunnumkal and Huseyin Topaloglu. A stochastic approximation algorithm for making pricing decisions in network revenue management problems. *Journal of Revenue and Pricing Management*, 9(5):419–442, 2010.

[96] Michael Landsberger and Isaac Meilijson. Intertemporal price discrimination and sales strategy under incomplete information. *The RAND Journal of Economics*, pages 424–430, 1985.

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

[98] Yanzhe Murray Lei, Stefanus Jasin, Joline Uichanco, and Andrew Vakhutinsky. Randomized product display (framing), pricing, and order fulfillment for e-commerce retailers. *Stefanus and Uichanco, Joline and Vakhutinsky, Andrew, Randomized Product*

Display (Framing), Pricing, and Order Fulfillment for E-commerce Retailers (November 9, 2018), 2018.

[99] Yuri Levin, Jeff McGill, and Mikhail Nediak. Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Production and Operations Management*, 19(1):40–60, 2010.

[100] Qian Liu and Garrett J Van Ryzin. Strategic capacity rationing to induce early purchases. *Management Science*, 54(6):1115–1131, 2008.

[101] Ilan Lobel. Dynamic pricing with heterogeneous patience levels. 2017.

[102] Brendan Lucier. An economic view of prophet inequalities. *ACM SIGecom Exchanges*, 16(1):26–49, 2017.

[103] Constantinos Maglaras and Joern Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management*, 8(2):136–148, 2006.

[104] Vahideh H Manshadi, Shayan Oveis Gharan, and Amin Saberi. Online stochastic matching: Online actions based on offline statistics. *Mathematics of Operations Research*, 37(4):559–573, 2012.

[105] Benny Mantin and Daniel Granot. Dynamic pricing in the presence of store-switching consumers. Technical report, Working paper, University of Waterloo, Canada, 2010.

[106] Dana Mattioli. On Orbitz, Mac Users Steered to Pricier Hotels. *The Wall Street Journal*, August 2012.

[107] Konrad Mierendorff. Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory*, 161:190–222, 2016.

[108] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

[109] Jamie Morgenstern and Tim Roughgarden. Learning simple auctions. In Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir, editors, *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 1298–1318. PMLR, 23–26 Jun 2016.

[110] Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[111] Peter J Olver. Chapter 20: Nonlinear ordinary differential equations. 2012.

[112] Nikolay Osadchiy and Gustavo Vulcano. Selling with binding reservations in the presence of strategic consumers. *Management Science*, 56(12):2173–2190, 2010.

[113] Mallesh M Pai and Rakesh Vohra. Optimal dynamic auctions and simple index rules. *Mathematics of Operations Research*, 38(4):682–697, 2013.

[114] Alessandro Pavan, Ilya Segal, and Juuso Toikka. Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2):601–653, 2014.

[115] Iosif Pinelis. Optimum bounds for the distributions of martingales in banach spaces. *The Annals of Probability*, pages 1679–1706, 1994.

[116] Martin I Reiman and Qiong Wang. An asymptotically optimal policy for a quantity-based network revenue management problem. *Mathematics of Operations Research*, 33(2):257–282, 2008.

[117] R Tyrrell Rockafellar. *Convex analysis*. Princeton university press, 1970.

[118] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.

[119] Aviad Rubinstein. Beyond matroids: Secretary problem and prophet inequality with general constraints. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 324–332. ACM, 2016.

[120] Aviad Rubinstein and Sahil Singla. Combinatorial prophet inequalities. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 1671–1687. Society for Industrial and Applied Mathematics, 2017.

[121] Uwe Saint-Mont. A simple derivation of a complicated prophet region. *Journal of Multivariate Analysis*, 80:67–72, 2002.

[122] Esther Samuel-Cahn. Comparisons of threshold stop rule and maximum for independent nonnegative random variables. *The Annals of Probability*, 12(4):1213–1216, 1983.

[123] Carl Shapiro and Hal R. Varian. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press, Cambridge, MA, 1999.

[124] Zuo-Jun Max Shen and Xuanming Su. Customer behavior modeling in revenue management and auctions: A review and new research opportunities. *Production and operations management*, 16(6):713–728, 2007.

[125] Artyom Shneyerov. An optimal slow dutch auction. *Economic Theory*, 57(3):577–602, 2014.

[126] Nancy L Stokey. Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371, 1979.

[127] Xuanming Su. Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5):726–741, 2007.

[128] Kalyan Talluri and Garrett Van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004.

[129] Kalyan T Talluri and Garrett J Van Ryzin. *The theory and practice of revenue management*, volume 68. Springer Science & Business Media, 2006.

[130] Alberto Vera and Siddhartha Banerjee. The bayesian prophet: A low-regret framework for online decision making. In *Abstracts of the 2019 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems*, pages 81–82. ACM, 2019.

[131] KW Wang. Optimum seat allocation for multi-leg flights with multiple fare types. In *AGIFORS PROCEEDINGS–*, 1983.

[132] Ruqu Wang. Optimal pricing strategy for durable-goods monopoly. *Journal of Economic Dynamics and Control*, 25(5):789–804, 2001.

[133] Shining Wu, Qian Liu, and Rachel Q Zhang. The reference effects on a retailer's dynamic pricing and inventory strategies with strategic consumers. *Operations Research*, 63(6):1320–1335, 2015.

[134] Qiqi Yan. Mechanism design via correlation gap. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete algorithms*, SODA '11, pages 710—719. ACM-SIAM, 2011.

[135] Hui Yang, Huaming Song, and Song Zhang. Dynamic pricing with strategic and myopic consumers. In *Business Intelligence and Financial Engineering (BIFE), 2010 Third International Conference on*, pages 104–108. IEEE, 2010.

[136] Chuiri Zhou and Yu Wu. Study on revenue management considering strategic customer behavior. *Journal of Service Science and Management*, 4(04):507, 2011.

Appendix A

Appendix to Chapter 1

A.1 Technical results

Proposition 10. *The function $\phi(t)$ is lower semi-continuous.*

Proof. We need to show that for all $t_0 \geq 0$, it holds that

$$\liminf_{t \to t_0} \phi(t) \geq \phi(t_0). \tag{A.1}$$

First, note that from the definition of $\phi(t)$, we are looking for a value v to set $\phi(t) = v$, i.e., v must be the smallest valuation verifying: $U(t, v) \geq U(t', v)$, for all $t' > t$—the inequality holds for all v setting $t' = t$ and therefore we can restrict the condition for t' strictly greater than t—, or equivalently,

$$\mathrm{e}^{-\mu t}(v - p(t)) \geq \mathrm{e}^{-\mu t'}(v - p(t')),$$

where by isolating v we get

$$v \geq \frac{p(t) - \mathrm{e}^{-\mu(t'-t)} p(t')}{1 - \mathrm{e}^{-\mu(t'-t)}}.$$

Observe that due to the lower semi-continuity of $p(t)$, the definition of the threshold function ϕ is equivalent to

$$\phi(t) = \sup_{t' > t} \left\{ \frac{p(t) - \mathrm{e}^{-\mu(t'-t)} p(t')}{1 - \mathrm{e}^{-\mu(t'-t)}} \right\}. \tag{A.2}$$

To prove that ϕ is lower semi-continuous we need the following auxiliary result, whose proof follows from the definition of liminf for functions.

Auxiliary lemma. If f and g are functions such that for all $y \geq x$ it holds that $f(x) \geq g(x,y)$, then $\liminf_{x\to x_0} f(x) \geq \liminf_{x\to x_0} g(x,y)$, for all $y \geq x_0$.

From (A.2) we have $\phi(t) \geq \frac{p(t)-\mathrm{e}^{-\mu(t'-t)}p(t')}{1-\mathrm{e}^{-\mu(t'-t)}}$ for all $t' \geq t$, and using the auxiliary lemma it follows that, for all $t' \geq t_0$,

$$\liminf_{t\to t_0} \phi(t) \geq \liminf_{t\to t_0} \frac{p(t)-\mathrm{e}^{-\mu(t'-t)}p(t')}{1-\mathrm{e}^{-\mu(t'-t)}}.$$

Due to the lower semi-continuity of $p(t)$ and the continuity of the exponential function, and in view of (A.1), the right side of this inequality is at least $\frac{p(t_0)-\mathrm{e}^{-\mu(t'-t_0)}p(t')}{1-\mathrm{e}^{-\mu(t'-t_0)}}$, and therefore,

$$\liminf_{t\to t_0} \phi(t) \geq \frac{p(t_0)-\mathrm{e}^{-\mu(t'-t_0)}p(t')}{1-\mathrm{e}^{-\mu(t'-t_0)}} \quad \forall t' \geq t_0.$$

Then, $\liminf_{t\to t_0} \phi(t)$ is at least the maximum, over all $t' \geq t_0$, of $\frac{p(t_0)-\mathrm{e}^{-\mu(t'-t_0)}p(t')}{1-\mathrm{e}^{-\mu(t'-t_0)}}$, which is equal to $\phi(t_0)$. Thus, ϕ is lower semi-continuous in $\mathbb{R}_0^+$. □

Lemma 10. *The Euler-Lagrange equation associated to the problem*

$$\max_p \int_0^{+\infty} G(t,p(t),p'(t),p''(t))\mathrm{d}t$$

is given by

$$f'\left(p(t)-\frac{p'(t)}{\mu}\right)\left(-\frac{p''(t)}{\mu}+p'(t)\right)(-\delta p(t)+p'(t))+f\left(p(t)-\frac{p'(t)}{\mu}\right)\left[\delta(\delta-\mu)p(t)-2\delta p'(t)+2p''(t)\right]=0. \tag{A.3}$$

Proof. Recall that $G(t,p(t),p'(t),p''(t)) = \mathrm{e}^{-\delta t}p(t)\left(-p'(t)+\frac{p''(t)}{\mu}\right)f\left(p(t)-\frac{p'(t)}{\mu}\right)$. We have to check that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial G}{\partial p''} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial p'} + \frac{\partial G}{\partial p} = 0 \tag{A.4}$$

is equivalent to equation (A.3).

The first term of the RHS of (A.4) is given by

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial G}{\partial p''} \;=\;& \frac{\mathrm{e}^{-\delta t}}{\mu}f\left(p(t)-\frac{p'(t)}{\mu}\right)\left(p''(t)-2\delta p'(t)+\delta^2 p(t)\right)+\\ & \frac{\mathrm{e}^{-\delta t}}{\mu}f'\left(p(t)-\frac{p'(t)}{\mu}\right)\left[2(p'(t)-\delta p(t))\left(p'(t)-\frac{p''(t)}{\mu}\right)+p(t)\left(p''(t)-\frac{p'''(t)}{\mu}\right)\right]+\\ & \frac{\mathrm{e}^{-\delta t}}{\mu}f''\left(p(t)-\frac{p'(t)}{\mu}\right)p(t)\left(p'(t)-\frac{p''(t)}{\mu}\right)^2.\end{aligned}$$

On the other hand, computing the second term we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial G}{\partial p'} &= \frac{\mathrm{e}^{-\delta t}}{\mu} f\left(p(t)-\frac{p'(t)}{\mu}\right)(\delta p(t)-p'(t))+\\ &\frac{\mathrm{e}^{-\delta t}}{\mu} f'\left(p(t)-\frac{p'(t)}{\mu}\right)\left[\left(\frac{p''(t)}{\mu}-p'(t)\right)\left(\frac{\delta p(t)-p'(t)}{\mu}+p(t)\right)+\frac{p(t)}{\mu}\left(p''(t)-\frac{p'''(t)}{\mu}\right)\right]+\\ &\frac{\mathrm{e}^{-\delta t}}{\mu} f''\left(p(t)-\frac{p'(t)}{\mu}\right)p(t)\left(\frac{p''(t)}{\mu}-p(t)\right)^2.\end{aligned}$$

Finally, the partial derivative of G with respect to p is the following

$$\frac{\partial G}{\partial p}=\mathrm{e}^{-\delta t}\left(\frac{p''(t)}{\mu}-p'(t)\right)\left(f\left(p(t)-\frac{p'(t)}{\mu}\right)+p(t)f'\left(p(t)-\frac{p'(t)}{\mu}\right)\right).$$

Thus, (A.3) comes from using the expressions above and equalizing the LHS of (A.4) to zero.

$\square$

A.2 Proofs for Section 1.5

A.2.1 Proof of Proposition 1

Let $p(t)$ be an optimal solution of the relaxed problem $[SPO_0^r]$ and suppose that there exists t such that $\psi(t)$ is an inner local maximum. Then, it must hold that

$$\psi'(t)=p'(t)-\frac{p''(t)}{\mu}=0. \tag{A.5}$$

Recalling that the valuation density f is positive, observe that, at t, the Euler-Lagrange equation (1.2) becomes

$$\delta(\delta-\mu)p(t)-2\delta p'(t)+2p''(t)=0,$$

and therefore, together with (A.5), $p'(t)=\frac{\delta p(t)}{2}$, and thus, $p'(t)-\delta p(t)<0$.

Let $\varepsilon>0$ and $\rho>0$ be such that $t_1=t-\varepsilon$ and $t_2=t+\rho$ satisfying $\psi(t_1)=\psi(t_2)$ and $p'(t_\mathrm{i})-\delta p(t_\mathrm{i})<0$ for $\mathrm{i}=1,2$. Furthermore, since ψ has a maximum at t, it must hold that $\psi'(t_1)>0$ and $\psi'(t_2)<0$.

Let us first suppose that $f'(\psi(t_1))=f'(\psi(t_2))\geq 0$. In this case, considering the first term in (1.2),

$$\underbrace{f'\left(p(t_2)-\frac{p'(t_2)}{\mu}\right)}_{f'(\psi(t_2))\geq 0}\underbrace{\left(-\frac{p''(t_2)}{\mu}+p'(t_2)\right)}_{\psi'(t_2)<0}\underbrace{(-\delta p(t_2)+p'(t_2))}_{<0}\geq 0,$$

and therefore, since $p(t)$ satisfies the Euler-Lagrange equation (1.2) for all t—and, in particular, for t_2—, we must have

$$\delta(\delta-\mu)p(t_2) - 2\delta p'(t_2) + 2p''(t_2) \leq 0. \tag{A.6}$$

Since by construction $p'(t_2) - \delta p(t_2) < 0$, we have $\frac{p'(t_2)}{\delta} < p(t_2)$, and bounding from below the first term in the LHS of (A.6), we obtain

$$-(\delta+\mu)p'(t_2) + 2p''(t_2) < 0. \tag{A.7}$$

Recalling that $t_2 = t + \rho$, taking the liminf in the LHS of (A.7) when $\rho \to 0$, by the lower semi-continuity of the price function $p(t)$, we obtain

$$-(\delta+\mu)p'(t) + 2p''(t) \leq 0,$$

which is equivalent to $2p''(t) \leq (\delta+\mu)p'(t)$. But from (A.5), $\mu p'(t) = p''(t)$, so $2\mu p'(t) \leq (\delta+\mu)p'(t)$ and therefore $\mu \leq \delta$ (because $p'(t) = \frac{\delta p(t)}{2} > 0$), which is a contradiction.

Now, consider the case where $f'(\psi(t_1)) = f'(\psi(t_2)) < 0$. Then, it must hold that

$$\underbrace{f'\left(p(t_1) - \frac{p'(t_1)}{\mu}\right)}_{f'(\psi(t_1))<0}\underbrace{\left(-\frac{p''(t_1)}{\mu} + p'(t_1)\right)}_{\psi'(t_1)>0}\underbrace{(-\delta p(t_1) + p'(t_1))}_{<0} > 0$$

and now we can proceed analogously to the argument above.

Therefore $\psi(t)$ cannot have an inner local maximum, and with a similar argument, neither an inner local minimum. Hence, $\psi(t)$ has to be monotone.

We are now left with showing that the function $\psi(t)$ is indeed non increasing. By contradiction, suppose that ψ is increasing. We will see that if so we could improve the expected revenue, contradicting that ψ corresponds to the optimal solution of the relaxed problem $[SPO_0^r]$. To this end, let us consider the constant function $\hat{p}(t) = p(0)$ for all t. Then, $\hat{\psi}(t) = p(0)$ and therefore the value of the objective function of $[SPO_0^r]$ by considering the feasible pricing policy $\hat{p}$ is given by

$$\hat{p}(0)(1 - F(\hat{\psi}(0))) + \int_0^\infty \mathrm{e}^{-\delta t}\hat{p}(t)(-\hat{\psi}(t))f(\hat{\psi}(t))\mathrm{d}t = p(0)(1 - F(p(0))).$$

On the other hand, the expected revenue of the seller under the pricing policy p can be computed as

$$p(0)(1 - F(\psi(0))) + \int_0^\infty \mathrm{e}^{-\delta t}p(t)(-\psi(t))f(\psi(t))\ \mathrm{d}t.$$

Note that the second term is negative and therefore the expression above is upper bounded by the expected revenue obtained by selling at time 0. That is,

$$p(0)(1-F(\psi(0))) + \int_0^\infty \mathrm{e}^{-\delta t} p(t)(-\psi(t)) f(\psi(t))\ \mathrm{d}t < p(0)(1-F(\psi(0))).$$

Note that $1-F(\psi(0)) > 1-F(p(0))$, and therefore the expected revenue under the price function $\hat{p}$ is greater than the expected revenue under the price function p, which contradicts the optimality of p. Thus, we can conclude that ψ is a non increasing function.

□

A.2.2 Proof of Theorem 1

Given a pair $(p(t), \psi(t))$ solution of $[SPO_0^r]$, with $\psi(t) = p(t) - \frac{p'(t)}{\mu}$ for all t, we must show that it meets the equilibrium constraint of $[SPO_0]$, that is:

$$t \in \arg\max_{s \geq 0}\ \mathrm{e}^{-\mu s}(\psi(t) - p(s)) \ \ \forall t. \tag{A.8}$$

Let $h(s) = \mathrm{e}^{-\mu s}(\psi(t) - p(s))$, leading to

$$h'(s) = \mathrm{e}^{-\mu s}(-\mu(\psi(t) - p(s)) - p'(s)),$$

and

$$h''(s) = -\mu \mathrm{e}^{-\mu s}(-\mu(\psi(t) - p(s)) - p'(s)) + \mathrm{e}^{-\mu s}(\mu p'(s) - p''(s)).$$

Given an interior solution t of (A.8), it must verify $h'(t) = 0$ and

$$h''(t) = \mu \mathrm{e}^{-\mu t}\left(p'(t) - \frac{p''(t)}{\mu}\right).$$

Since $(p(t), \psi(t))$ is solution of $[SPO_0^r]$, then from Proposition 1 we know that $\psi'(t) \leq 0$, and therefore, $h''(t) \leq 0$. Hence, $t \in \arg\max_{s\geq 0}\ \mathrm{e}^{-\mu s}(\psi(t) - p(s))$, for any pair of functions $(p(t), \psi(t))$ solution of $[SPO_0^r]$. Recalling that the solution of $[SPO_0^r]$ defines an upper bound of $[SPO_0]$, we have that such pair $(p(t), \psi(t))$ indeed defines a solution to $[SPO_0]$.

□

A.2.3 Proof of Lemma 1

By contradiction, suppose that for $T = \ln(1/c)/\delta$, we have that:

$$\int_T^\infty \mathrm{e}^{-\delta t} p(t) \mathrm{d}(1-F(\psi(t))) > c\left[p(0)(1-F(\psi(0))) + \int_0^\infty \mathrm{e}^{-\delta t} p(t) \mathrm{d}(1-F(\psi(t)))\right]. \tag{A.9}$$

Consider the price function $\hat{p}(t) = p(t+T)$ and its associated purchasing function $\hat{\psi}$. The seller's expected revenue can be computed as:

$$R_{\hat{p}} = \hat{p}(0)(1 - F(\hat{\psi}(0))) + \int_0^\infty \mathrm{e}^{-\delta t}\hat{p}(t)\mathrm{d}(1 - F(\hat{\psi}(t))).$$

By the definition of $\hat{p}$ and doing the change of variable $u = t + T$, it follows that the seller's expected revenue is given by:

$$R_{\hat{p}} = p(T)(1 - F(\psi(T))) + \mathrm{e}^{\delta T} \int_T^\infty p(t)\mathrm{e}^{-\delta t}\mathrm{d}(1 - F(\psi(t))).$$

Applying (A.9), it follows that this expression verifies

$$R_{\hat{p}} > p(T)(1 - F(\psi(T))) + \mathrm{e}^{\delta T} c \left[p(0)(1 - F(\psi(0))) + \int_0^\infty \mathrm{e}^{-\delta t} p(t)\mathrm{d}(1 - F(\psi(t))) \right].$$

Note that $p(T)(1 - F(\psi(T)))$ is non negative, and that $T = \ln(1/c)/\delta$ implies $\mathrm{e}^{\delta T} c = 1$. Thus, the seller's expected revenue for the pricing policy $\hat{p}$ is bigger than the seller's expected revenue for the pricing policy p, which contradicts the optimality of the price function p.

□

A.3 Proof for Section 1.6

A.3.1 Proof of Lemma 2

Without loss of generality let us suppose that $k = 1$, that is, the buyer arrives at time $t_0 + t$ belonging to $\tilde{I}_1$ with valuation $v \geq p(T)$, and further assume that he will not purchase before time $T + t_0$.

To prove the lemma we analyze the consumer behaviour in the unobservable case under the pricing policy $\hat{p}$ depending on his valuation. More specifically we will prove the followings three statements:

1. If $v \in [p(T), \psi(T))$, then the buyer buys at time $2T + t_0$.
2. If $v \in [\psi(T), \psi(0))$, then the buyer waits and buys at time $\tau \in (T+t_0, 2T+t_0]$ satisfying $\psi(\tau) = v$.
3. If $v \geq \psi(0)$ the buyer purchases at time $t_0 + T$.

First, consider a buyer with valuation $v \in [p(T), \psi(T))$. Knowing that he will purchase to gain some positive utility (eventually at time $2T+t_0$), if he decides to buy at time $\tau < 2T+t_0$, then by the monotonicity of the purchasing function ψ in the observable case, we have that $\psi(\tau - (T+t_0)) > \psi(2T+t_0-(T+t_0)) = \psi(T)$ and it means that the buyer must have valuation greater than $\psi(T)$ to be optimum to purchase at time τ, which is not the case. We then conclude that in this case he will buy at time $2T+t_0$.

Secondly, if the buyer has valuation $v \in [\psi(T), \psi(0))$, then by using the calculation of the purchasing function for the observable arrival cas—conducted under the assumption that the buyer arrives at time 0—, we have that for some $t \in [0, T]$, it holds that $v = \psi(t)$, i.e.,

$$t \in \arg\max_{s \geq 0} U(s, \psi(t)),$$

which means that

$$\mathrm{e}^{-\mu t}(\psi(t) - p(t)) \geq \mathrm{e}^{-\mu s}(\psi(t) - p(s)), \ \forall s \geq 0.$$

This is equivalent to

$$\mathrm{e}^{-\mu(T+t_0)}\mathrm{e}^{-\mu t}(\psi(t) - p(t)) \geq \mathrm{e}^{-\mu(T+t_0)}\mathrm{e}^{-\mu s}(\psi(t) - p(s)), \ \forall s \geq 0.$$

Hence, the buyer will buy at time $\tau = T + t_0 + t$ satisfying $\psi(t) = v$.

Finally, the third statement follows directly from the definition of the threshold function ψ.

The lemma follows by observing that if the buyer has valuation at least $\psi(T)$, the seller's revenue is the same as in the observable case with the buyer arriving at time $T + t_0$ and accumulating revenue up to time $2T + t_0$ (cases (2) and (3)). But if the buyer has valuation between $p(T)$ and $\psi(T)$ (case (1)), then he will buy before time $2T + t_0$ in the unobservable setting under the price function $\hat{p}$ but he will buy after that time in the observable case with arrival time $T + t_0$.

Therefore, we conclude that, conditioned on the event that the buyer with valuation greater than $p(T)$ arrives at time $T + t_0$—which is equivalent to looking at the problem in the interval $[T+t_0, 2T+t_0]$ in the observable case—, the seller's expected revenue under the policy $\hat{p}$ in the unobservable case is at least the expected revenue earned up to time $2T + t_0$ in the observable case with arrival time $T + t_0$.

A.4 Proof for Section 1.7.1

A.4.1 Proof of Theorem 2

We will compute a lower bound of the seller's expected revenue for the unobservable case. For that purpose, consider the pricing policy $\hat{p}$ described in Figure 1.4 and fix the buyer arrival time τ. Recall that t_0 is the uniform random variable involved in the random shift applied over the original price function p to get $\hat{p}$. By defining $T = \ln(1/c)/\delta$, the price function has period $2T$.

Suppose, without loss of generality, that the buyer arrives during the first period; i.e., $\tau \in [t_0, t_0 + 2T]$. Thus, $t_0 \sim \text{Unif}[\tau - 2T, \tau]$. In order to have intervals defined around t_0, we denote $\tilde{\mathcal{I}}_1 := [\tau - T, \tau]$ and $\tilde{\mathcal{I}}_2 := [\tau - 2T, \tau - T]$. With this definition, we have that $\tau \in \tilde{I}_{\mathrm{i}}$ if and only if $t_0 \in \tilde{\mathcal{I}}_{\mathrm{i}}$, for $\mathrm{i} = 1, 2$.

Let us denote by R_τ^{uo} the seller's revenue in the unobservable case if the arrival time is τ. We only consider the buyer's arrival if it belongs to the interval $\tilde{I}_1$, otherwise, we simply bound the revenue by 0.

Note that if $\tau \in \tilde{I}_1$, we can lower bound the expected value of R_τ^{uo} by the expected revenue obtained by considering that the buyer has valuation at least $p(T)$ and that he purchases after time $t_0 + T$. This is because the buyer does not purchase if $v < p(T)$, and by waiting up to $t_0 + T$ to buy when he could buy would hurt the seller's revenue given her discount factor.

Then, by Lemma 2, $\mathbb{E}(R_\tau^{uo})$ is at least the expected revenue earned up to time $2T + t_0$ in the observable case with arrival time $T + t_0$. Applying Lemma 1, we have $\mathbb{E}(R_\tau^{uo}) \geq (1-c)R_{t_0+T}$, where R_{t_0+T} denotes the expected revenue in the observable case if the buyer arrives at time $t_0 + T$.

We now use the analysis above to compute a bound for the expected value of the seller's revenue in the unobservable case conditioned on the event that the buyer arrives at time τ.

$$\begin{aligned}
\mathbb{E}(R_\tau^{uo}) &= \mathbb{E}_{t_0}(\mathbb{E}(R_\tau^{uo} \mid t_0)) \\
&= \mathbb{E}(R_\tau^{uo} \mid t_0 \in \mathcal{I}_1)\mathbb{P}(t_0 \in \mathcal{I}_1) + \mathbb{E}(R_\tau^{uo} \mid t_0 \in \mathcal{I}_2)\mathbb{P}(t_0 \in \mathcal{I}_2) \\
&= \frac{1}{2}\mathbb{E}(R_\tau^{uo} \mid t_0 \in \mathcal{I}_1) + \frac{1}{2}\mathbb{E}(R_\tau^{uo} \mid t_0 \in \mathcal{I}_2) \\
&\geq \frac{1}{2}(1-c)\mathbb{E}_{t_0}(R_{t_0+T} \mid t_0 \in \mathcal{I}_1),
\end{aligned}$$

where the last equality holds because $t_0 \sim \text{Unif}[\tau - 2T, \tau]$ and the inequality follows from the analysis above. Note that $R_{t_0+T} = c\mathrm{e}^{-\delta(t_0-\tau)}R_\tau$, with $\mathrm{e}^{-\delta T} = c$, and therefore it is enough to compute $\mathbb{E}_{t_0}\left(\mathrm{e}^{-\delta(t_0-\tau)} \mid t_0 \in \mathcal{I}_1\right)$. In fact,

$$\begin{aligned}
\mathbb{E}_{t_0}\left(\mathrm{e}^{-\delta(t_0-\tau)} \mid t_0 \in \mathcal{I}_1\right) &= \int_{\tau-T}^{\tau} \mathrm{e}^{-\delta(t_0-\tau)}\frac{1}{T}\mathrm{d}t_0 \\
&= \frac{\mathrm{e}^{\delta T} - 1}{\delta T}.
\end{aligned}$$

By the definition of T, we know that $T\delta = \ln(1/c)$ and $\mathrm{e}^{\delta T} = 1/c$, and therefore we have

$$\mathbb{E}_{t_0}\left(\mathrm{e}^{-\delta(t_0-\tau)} \mid t_0 \in \mathcal{I}_1\right) = \frac{1-c}{c\ln(1/c)}.$$

We then obtain the following lower bound for the expectation of the seller's revenue in the unobservable case that depends on c:

$$\mathbb{E}(R_\tau^{uo}) \geq \frac{(1-c)^2}{2\ln(1/c)} R_\tau.$$

Noting that R_τ is the expected value of the seller's revenue in the observable case with buyer's time arrival τ, follows that for each time arrival τ, the ratio between the expected revenue in the observable and the unobservable case is at most

$$\frac{\mathbb{E}(\text{Rev. Obs} \mid \tau)}{\mathbb{E}(\text{Rev. Unobs} \mid \tau)} \leq \frac{2\ln(1/c)}{(1-c)^2}.$$

The latter expression is minimized at $c = \mathrm{e}^{W_{-1}(-1/(2\sqrt{\mathrm{e}}))+1/2} \approx 0.284$ and the minimum is $-\frac{2W_{-1}(-1/(2\sqrt{\mathrm{e}}))+1}{\left(\mathrm{e}^{W_{-1}(-1/(2\sqrt{\mathrm{e}}))+1/2}-1\right)^2}$, which is roughly 4.911.

$\square$

Appendix B

Appendix to Chapter 2

B.1 Proofs for Section 2.4

B.1.1 Proof of Proposition 2

Let us first show that $F(x) \geq u$ implies $F^{-1}(u) \leq x$. In fact, if $F(x) \geq u$ then $F^{-1}(u) = \inf\{w : F(w) \geq u\} \leq x$, and we obtain the desire inequality.

It remains to prove that $F^{-1}(u) \leq x$ implies $F(x) \geq u$. If $F^{-1}(u) \leq x$, then $x \geq F^{-1}(u) = \inf A = x_0$, where $A = \{w : F(w) \geq u\}$. We divide the rest of the proof in two cases depending on whether x_0 belongs to A or not.

Case 1: $x_0 \in A$. In this case, $F(x_0) \geq u$ and from the monotonicity of F together with $x \geq x_0$, follows that $F(x) \geq u$.

Case 2: $x_0 \notin A$. In this case, there exists a sequence $\{x_n\}_n \subset A$ such that $x_n \to x_0$ and $x_{\mathrm{i}} > x_0$ $\forall$i. F is a distribution and by its right-continuity we have that $F(x_0) = \lim_{n\to\infty} F(x_{\mathrm{i}}) \geq u$. Again, from the monotonicity of F together with $x \geq x_0$, we conclude that $F(x) \geq F(x_0) \geq u$, and we conclude the proof.

$\square$

B.1.2 Proof of Proposition 3

Let us call W to the random variable $F^{-1}(U)$. We will prove that W is distributed according to F. In effect,

$$\mathbb{P}(W \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x),$$

where the first equality follows from the definition of W, the second equality holds due to Proposition 2, and the last one because U is uniformly distributed. We conclude that $F^{-1}(U) \sim F$. $\square$

B.2 Proof for Section 2.5

B.2.1 Proof of Lemma 3

Let us first assume that G is non-regular. Recall from [110] that $\bar{\phi}$ is constructed trough *ironing* of the virtual valuation function as follows. Take $Q(\theta) = \theta G^{-1}(1-\theta)$ and $R(\cdot)$ the concave hull of $Q(\cdot)$, i.e.,

$$R(\theta) = \min \left\{ xQ(\theta_1) + (1-x)Q(\theta_2) : x\theta_1 + (1-x)\theta_2 = \theta \text{ and } x, \theta_1, \theta_2 \in [0,1] \right\}.$$

That is, R is the smallest concave function on $[0,1]$ that is above Q. We now define the ironed virtual value as $\bar{\phi}(v) = R'(1 - G(v))$ and we denote by $\bar{\phi}^{-1}$ its generalized inverse. Thus,

$$\mathbb{E}(\bar{\phi}(v) \mid \bar{\phi}(v) \geq \tau)\mathbb{P}(\bar{\phi}(v) \geq \tau) = \int_{\bar{\phi}^{-1}(\tau)}^{\infty} \bar{\phi}(u)\mathrm{d}G(u) = \int_0^q R'(\theta)\mathrm{d}\theta = R(q),$$

where the second equality follows by the definition of $\bar{\phi}$ and q together from performing the change of variable $\theta = 1 - G(u)$. Then, calling q_1, q_2 and x to the values such that $R(q) = xQ(q_1) + (1-x)Q(q_2)$ and using that $Q(\theta) = \theta G^{-1}(1-\theta)$, we obtain

$$R(q) = xq_1G^{-1}(1-q_1) + (1-x)q_2G^{-1}(1-q_2),$$

and the second part of the lemma follows.

If the distribution is regular, ϕ is non-decreasing implying that Q is already a concave function. Therefore, $R = Q$ and the result follows.

$\square$

B.2.2 Proof of Theorem 4

We denote by $\mathbb{P}_{t-1}$ the probability distribution conditional on the history $\mathcal{H}_{t-1}$, and the notation extends to the expectation.

We will prove prove that, given an instance for the multi-item auction problem, there exists an α-approximation mechanism. To this end, we define the algorithm and first compute its expected revenue. Then, we use that the elements selected by the auction are the same as the ones selected by the online selection algorithm over the instances taken. And finally, by Theorem 3 together with the approximation factor for the online selection algorithm, we obtain the result.

Note that the expected revenue of the mechanism is given by the sum, over the elements, of the price times the probability of selling the item.

To simplify the proof, assume first that the valuations are regular. In that case, we consider the mechanism given by Algorithm 1. Then, the prices obtained are $\left(\phi_{\mathrm{i}}^{+}\right)^{-1}(\tau_{\mathrm{i}})$, where τ_{i} is the threshold given by the online selection algorithm, and the expected revenue of the mechanism is given by

$$\sum_{t=1}^{n}\left(\phi_{\sigma(t)}^{+}\right)^{-1}(\tau_{\sigma(t)})\mathbb{P}_{t-1}\left(v_{\sigma(t)} \geq \left(\phi_{\sigma(t)}^{+}\right)^{-1}(\tau_{\sigma(t)})\right).$$

By Lemma 3 we have that the above summation equals to

$$\sum_{t=1}^{n}\mathbb{E}_{t-1}\left(\phi_{\sigma(t)}^{+}\middle|\phi_{\sigma(t)}^{+} \geq \tau_{\sigma(t)}\right)\mathbb{P}_{t-1}\left(\phi_{\sigma(t)}^{+}(v_{\sigma(t)}) \geq \tau_{\sigma(t)}\right).$$

Let χ_t be the indicator function of the event $\sigma(t) \in \mathrm{Alg}$, that is, $\chi_t = 1$ whenever $\sigma(t)$ is selected by the algorithm Alg. We can rewrite the expression above as

$$\mathbb{E}\left(\sum_{t=1}^{n}\phi_{\sigma(t)}^{+}(v_{\sigma(t)})\chi_t\right) = \mathbb{E}\left(\sum_{\mathrm{i}\in\mathrm{Mech}}\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}).\right)$$

Observe that since $v_{\mathrm{i}} \geq (\phi_{\mathrm{i}}^{+})^{-1}(\tau_{\mathrm{i}})$ if and only if $\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}}) \geq \tau_{\mathrm{i}}$ for all $\mathrm{i} \in \mathcal{I}$, Mech selects the same elements Alg does over the instance $(\mathcal{I}, \mathcal{G}^{\phi}, \mathcal{T})$. Therefore, as Alg is an α-approximation, it follows that

$$\mathbb{E}\left(\sum_{\mathrm{i}\in\mathrm{Mech}}\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}})\right) \geq \alpha\cdot\mathbb{E}\left(\max_{S\in\mathcal{T}}\sum_{\mathrm{i}\in S}\phi_{\mathrm{i}}^{+}(v_{\mathrm{i}})\right).$$

By Theorem 3, the latter is equal to the expected revenue of the optimal mechanism. We conclude that mechanism 1 is an α-approximation.

If a distribution is not regular, the posted price mechanism needs to randomize between two prices. Specifically, given $q_{\mathrm{i}} = \mathbb{P}(\phi_{\mathrm{i}}^{+}(v) \leq \tau_{\mathrm{i}})$ we consider the values $x, q_{\mathrm{i}}^1, q_{\mathrm{i}}^2 \in [0,1]$ obtained from Lemma 3, and define prices $p_{\mathrm{i}}^1 = G_{\mathrm{i}}^{-1}(1-p_{\mathrm{i}}^1)$ and $p_{\mathrm{i}}^2 = G_{\mathrm{i}}^{-1}(1-p_{\mathrm{i}}^2)$. Then the mechanism set price p_{i}^1 with probability $xq_{\mathrm{i}}^1/q_{\mathrm{i}}$ and price p_{i}^2 with probability $xq_{\mathrm{i}}^2/q_{\mathrm{i}}$. Then, the expected price is

$$\frac{xq_{\mathrm{i}}^1G_{\mathrm{i}}^{-1}(1-q_{\mathrm{i}}^1) + (1-x)q_{\mathrm{i}}^2G_{\mathrm{i}}^{-1}(1-q_{\mathrm{i}}^2)}{q_{\mathrm{i}}},$$

and the rest of the proof is analogous to the previous case by using Lemma 3.

□

B.3 Proofs for Section 2.6

B.3.1 Proof of Proposition 4

Let us first show the continuity of the function H. For $q \in [0,1)$, $H(q) = I(q)/L(q)$, where $I(q) = \int\limits_q^1 F^{-1}(y)\mathrm{d}y$ and $L(q) = 1-q$. In the interval $(0,1)$, the function $I(q) = \int_q^1 F^{-1}(y)\mathrm{d}y$ is differentiable due to the calculus fundamental theorem and $L(q) = 1-q$ is clearly differentiable and non-zero. Therefore, H is differentiable on $(0,1)$, and in particular continuous.

To study the monotonicity of H let us compute the first derivative of H and analyze its sign. Observe that

$$H'(q) = \frac{-F^{-1}(q)(1-q) + \int_q^1 F^{-1}(y)\mathrm{d}y}{(1-q)^2} = \frac{1}{(1-q)^2}\int\limits_q^1 (F^{-1}(y) - F^{-1}(q))\mathrm{d}y,$$

and $F^{-1}(y) - F^{-1}(q) \geq 0$, since F^{-1} is non-decreasing, and y is at least q. Therefore, $H'(q) = 0$ if and only if F^{-1} equals $F^{-1}(q)$ almost everywhere in $[q,1]$, which in turn happens if and only if $\lim\limits_{s \to F^{-1}(q)^-} F(s) = q$ and $F(F^{-1}(q)) = 1$. Taking $T = q$ the proof follows.

$\square$

B.3.2 Proof of Proposition 5

1. Note that G is non-decreasing and right-continuous due to Proposition 4. Furthermore, from the definition follows that $\lim_{t\to-\infty} G(t) = 0$ and $\lim_{t\to\infty} G(t) = 1$, and thus G is a distribution. The support of G comes from the fact that $H(0) = \int_0^1 F^{-1}(y)\mathrm{d}y = \mathbb{E}(w)$, $H^{-1}(1) = \omega_1(F)$, and G is strictly increasing in the interior of the interval.

2. Let t be in $[\mathbb{E}(w), \omega_1(F))$. By Proposition 4, G is strictly increasing and continuous on this interval, and therefore invertible. It is then sufficient to show that $\phi_G(H(q)) = F^{-1}(q)$ where $H(q) = t$. In particular, $q \in [0,T)$, with T as in the statement of Proposition 4. Since G is also differentiable and $G^{-1} = H$ in this interval, it follows that
$$\phi_G(H(q)) = H(q) - \frac{1-q}{G'(H(q))} = H(q) - (1-q)H'(q).$$
On the other hand, from the definition of H we have that
$$H'(q) = \frac{-F^{-1}(q)(1-q) + \int_q^1 F^{-1}(y)\mathrm{d}y}{(1-q)^2} = \frac{-F^{-1}(q) + H(q)}{1-q},$$
and therefore $H(q) - (1-q)H'(q) = F^{-1}(q)$, which completes the proof.

3. It follows directly from (2) that ϕ_G is non-decreasing, as both F^{-1} and G are non-decreasing. Since $\phi_G(\mathbb{E}(w)) = F^{-1}(0) = 0$, it holds that ϕ_G is non-negative which in turn implies that $\phi_G^+ = \phi_G$.

B.3.3 Proof of Lemma 4

Recall that w is a random variable distributed according to F. Let U be a random variable uniformly distributed between 0 and 1, and define the random variable v as $v = G^{-1}(U)$, where G is given by the distribution described above. Applying Proposition 3 holds that v has distribution G and by Proposition 5 (3), $\phi_G^+ = \phi_G$ and therefore it remains to study the distribution of ϕ_G. By Proposition 5 (2), we observe that for $t < \omega_1(F)$,

$$\mathbb{P}(\phi_G(v) \le t) = \mathbb{P}(F^{-1}(G(v)) \le t) = \mathbb{P}(v \le G^{-1}(F(t))) = G(G^{-1}(F(t))) = F(t),$$

where the second equality follows from Proposition 2, the third holds because $v \sim G$, and the last equality follows since by Proposition 4, G is invertible in $[\mathbb{E}(w), \omega_1(F))$ and thus $G \circ G^{-1} = I$.

□

B.3.4 Proof of Theorem 5

Let Q_t be the event that item $\sigma(t)$ is selected by Algorithm 2. Recall that we denote by $\mathbb{P}_{t-1}$ the probability distribution conditional on the history $\mathcal{H}_{t-1}$, and the notation extends to the expectation. We denote by $\mathbb{1}_{Q_t}$ the indicator function of the event Q_t. By conditioning on the history, we have that

$$\mathbb{E}\left(\sum_{x \in \mathrm{Alg}} w_x\right) = \sum_{t=1}^{n} \mathbb{E}\left(w_{\sigma(t)} \mathbb{1}_{Q_t}\right) = \sum_{t=1}^{n} \mathbb{E}\left(\mathbb{E}_{t-1}\left(w_{\sigma(t)} \mathbb{1}_{Q_t}\right)\right).$$

For $t \in [n]$, let $x = \sigma(t)$ and $p_x = \mathcal{M}(\mathcal{H}_{t-1}, \mathcal{G}, x)$ be the price computed by $\mathcal{M}$. Setting $R(t) = t(1 - G_x(t))$, where G_x is the distribution of v_x, we claim that Algorithm 2 satisfies that

$$\mathbb{E}_{t-1}\left(w_x \mathbb{1}_{Q_t}\right) = R(p_x).$$

Before proving this, we see how to conclude the theorem using the equality above. Since $\mathcal{M}$ is an α-approximation and using Theorem 3, we have that

$$\sum_{x \in X} p_x(1 - G(p_x)) \ge \alpha \cdot \mathbb{E}\left(\max_{A \in \mathcal{T}} \sum_{x \in A} \phi_x^+(v_x)\right) = \alpha \cdot \mathbb{E}\left(\max_{A \in \mathcal{T}} \sum_{x \in A} w_x\right),$$

where in the last equality we used the fact that the valuations are obtained from the Valuation Mapping Lemma, and that the distributions in $\mathcal{F}$ are independent. We remark that for the last equality to be true it is sufficient that the distributions in $\mathcal{F}$ are independent, and that $\phi_x^+(v_x) \sim F_x$ for all $x \in X$. In general, if we remove the independence assumption then the equality does not hold. This proves that Algorithm 2 is an α-approximation.

It remains to prove the equality $\mathbb{E}_{t-1}(w_x \mathbb{1}_{Q_t}) = p_x(1-G_x(p_x))$. To this end, we condition on whether we are in line 4 or line 5 of Algorithm 2. If the condition in line 4 holds, then $\phi_x(v_x) > \phi_x(p_x)$ if and only if $v_x > p_x$. In particular, $\mathbb{P}_{t-1}(Q_t) = \mathbb{P}_{t-1}(\phi_x(v_x) > \phi_x(p_x)) = \mathbb{P}_{t-1}(v_x > p_x) = 1 - G_x(p_x)$. By Lemma 3 and Proposition 5,

$$\mathbb{E}_{t-1}\left(w_{\sigma(t)} | Q_t\right) = \mathbb{E}_{t-1}(\phi_x(v_x) | \phi_x(v_x) > \phi_x(p_x)) = p_x\,,$$

and putting all together we conclude that

$$\mathbb{E}_{t-1}\left(w_{\sigma(t)} \mathbb{1}_{Q_t}\right) = \mathbb{E}_{t-1}\left(w_{\sigma(t)} | Q_t\right) \mathbb{P}_{t-1}(Q_t) = p_x(1 - G_x(p_x)) = R(p_x).$$

Suppose now that the condition of line 5 is satisfied. By Proposition 6, the function F^{-1} is constant in the interval $[G(p_x^-), G(p_x^+))$. In fact, we can find an explicit expression for G_x in the interval $[p_x^-, p_x^+)$.

Given $q \in [G_x(p_x^-), G_x(p_x^+)]$, note that we have

$$\begin{aligned} H(q) = \frac{1}{1-q}\int_q^1 F^{-1}(y)\mathrm{d}y &= \frac{1}{1-q}\left[\int_q^{G_x(p_x^+)} F^{-1}(y)\mathrm{d}y + \int_{G_x(p_x^+)}^1 F^{-1}(y)\mathrm{d}y\right] \\ &= \frac{G_x(p_x^+) - q}{1-q}\phi_x(p_x) + \frac{1}{1-q}\int_{G_x(p_x^+)}^1 F^{-1}(y)\mathrm{d}y, \end{aligned}$$

where the last equality follows from F^{-1} being equal to $\phi_x(p_x)$ in the interval $[G_x(p_x^-), G_x(p_x^+))$. Then, for every $p \in [p_x^-, p_x^+)$, we have that

$$G_x(p) = \frac{p - \phi_x(p_x^+)G_x(p_x^+) - \int_{G(p_x^+)}^1 F^{-1}(y)\mathrm{d}y}{p - \phi_x(p_x^+)}.$$

Using the expression of G_x shown above, it follows that

$$R(p) = \phi_G(p_x^+)[G_x(p_x^+) - G_x(p)] + \int_{G_x(p_x^+)}^1 F^{-1}(y)\mathrm{d}y.$$

By the definition of θ_x, $G_x(p) = \theta_x G_x(p_x^-) + (1-\theta_x)G_x(p_x^+)$ and therefore $R(p) = \theta_x R(p_x^-) + (1-\theta_x)R(p_x^+)$. By conditioning on whether line 6 or line 7 holds, we have

$$\begin{aligned} \mathbb{E}_{t-1}\left(w_{\sigma(t)}\chi(Q_t)\right) &= \theta_x \mathbb{E}_{t-1}(\phi_x(v_x) | \phi_x(v_x) \geq \phi_x(p_x))\mathbb{P}_{t-1}(\phi_x(v_x) \geq \phi_x(p_x)) \\ &\quad + (1-\theta_x)\mathbb{E}_{t-1}(\phi_x(v_x) | \phi_x(v_x) > \phi_x(p_x))\mathbb{P}_{t-1}(\phi_x(v_x) > \phi_x(p_x)). \end{aligned}$$

Note that $\phi_x(v_x) \geq \phi_x(p_x)$ if and only if $v_x \geq p_x^-$. On the contrary, $\phi_x(v_x) > \phi_x(p_x)$ if and only if $v_x \geq p_x^+$. Thus, the expression above is equivalent to

$$\theta_x \mathbb{E}_{t-1}(\phi_x(v_x)|v_x \geq p_x^-)\mathbb{P}_{t-1}(v_x \geq p_x^-) + (1-\theta_x)\mathbb{E}_{t-1}(\phi_x(v_x)|v_x > p_x^+)\mathbb{P}_{t-1}(v_x > p_x^+).$$

Using Lemma 3 together with the definition of R, we conclude that

$$\begin{aligned}\mathbb{E}_{t-1}\left(w_{\sigma(t)}\mathbb{1}_{Q_t}\right) &= \theta_x p_x^-(1-G_x(p_x^-)) + (1-\theta_x)p_x^+(1-G_x(p_x^+)) \\ &= \theta_x R(p_x^-) + (1-\theta_x)R(p_x^+) = R(p_x).\end{aligned}$$

$\square$

Appendix C

Appendix to Chapter 3

C.1 Additional material

C.1.1 Static Probabilistic Control Heuristic

The *static probabilistic control heuristic* consists on setting the policy $\pi_t = \phi_1^*$ for all t until stock is out, where ϕ_1^* is an optimal solution of (3.2) taking $\rho = C/T$. This heuristic is described in Algorithm 4 and gave us a feasible solution of problem (3.1), namely J^S.

This heuristic was already studied in the revenue management literature (see e.g. [116], [34]) and its well know that the policy obtained from the SPC heuristic has a revenue loss of $O(\sqrt{T})$ for different online section problems. For completeness, we give a proof of this result, for the class of DRCRC problems, stated in the following lemma.

Lemma 11. *If we call J^S the value of the feasible solution of Problem (3.1) given by the SPC heuristic described above, the revenue loss is $O(\sqrt{T})$, that is,*

$$J^* - J^S \leq O(\sqrt{T}).$$

Proof. Note that taken an optimal solution of (3.2) for $\rho = C/T$, namely ϕ^*, we have that

$$J(C/T) = \sum_{\theta \in \Theta} p_\theta \int\limits_{\mathcal{A}} \mathbb{E}_\varepsilon \left(r(\theta, a^{\phi^*}, \varepsilon) \right) \mathrm{d}\phi_\theta^*(a).$$

The value of the objective function of problem (3.1) by considering the policy π that at each time period applies the optimal solution ϕ^* is given by

$$\sum_{t=1}^{T} \mathbb{E}_{\varepsilon_t} \left(\sum_{\theta_t \in \Theta} p_{\theta_t} \int_{\mathcal{A}} r(\theta_t, a_t^{\pi}, \varepsilon_t) \, \mathrm{d}\phi_{\theta}^*(a) \right) = TJ(C/T). \tag{C.1}$$

Since π might not be feasible, to get J^S we have to subtract from (C.1) the revenue associated to customers who are not served because of the lack of capacity, which is upper bounded by

$$\sum_{l=1}^{L} r_{\max}^{l} \mathbb{E} \left(\left(\sum_{t=1}^{T} y_l(\theta_t, a_t^{\pi}, \varepsilon_t) - C_l \right)^{+} \right),$$

where $r_{\max}^{l}$ is the largest extra revenue gain by increasing the capacity of resource l by one unit.

Thus,

$$J^* - J^S \leq \sum_{l=1}^{L} r_{\max}^{l} \mathbb{E} \left(\left(\sum_{t=1}^{T} y_l(\theta_t, a_t^{\pi}, \varepsilon_t) - C_l \right)^{+} \right)$$

and in the rest of the proof we will bound right hand side in the inequality above.

In what follows, we will write $y_{t,l}$ to refer to $y_l(\theta_t, a_t^{\pi}, \varepsilon_t)$.

Due to ϕ^* is an optimal solution of (3.2) for $\rho = C/T$, if the associated resource consumption for each resource l is given by y_l, we have that

$$T \sum_{\theta \in \Theta} \int_{\mathcal{A}} \mathbb{E}_{\varepsilon}(y_l) \, \mathrm{d}\phi_{\theta}^*(a) \leq C_l,$$

and therefore

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{t=1}^{T} y_{t,l}(\theta_t, a_t^{\pi}, \varepsilon_t) - C_l \right)^{+} \right) &\leq \mathbb{E} \left(\left(\sum_{t=1}^{T} y_{t,l} - T\mathbb{E}(y_l) \right)^{+} \right) \\ &\leq \mathbb{E} \left(\left| \sum_{t=1}^{T} y_{t,l} - T\mathbb{E}(y_l) \right| \right), \end{aligned}$$

where the last inequality follows because the positive part of a real value is always upper bounded by its absolute value.

Furthermore, the square root function is concave and applying Jensen's inequality together with the independence of variables $\{y_{t,l}\}_t$ and that $\mathbb{E}\left(\sum_{t=1}^{T} y_{t,l}\right) = T\mathbb{E}(y_l)$, we obtain

$$
\begin{aligned}
\mathbb{E}\left(\left|\sum_{t=1}^{T} y_{t,l} - T\mathbb{E}(y_l)\right|\right) &\leq \sqrt{\mathbb{E}\left(\left(\sum_{t=1}^{T} y_{t,l} - T\mathbb{E}\left(y_l\right)\right)^2\right)} \\
&= \sqrt{\operatorname{Var}\left(\sum_{t=1}^{T} y_{t,l}\right)} \\
&= \sqrt{\sum_{t=1}^{T} \operatorname{Var}\left(y_{t,l}\right)}\,.
\end{aligned}
$$

Note that $\operatorname{Var}(y_{t,l}) \leq \mathbb{E}\left(y_{t,l}^2\right)$, and by assumption made at the beginning of Section 3.6 we obtain

$$
\sqrt{\sum_{t=1}^{T} \operatorname{Var}\left(y_{t,l}\right)} \leq \sqrt{T}\bar{y}_{\infty}\,.
$$

Putting all together, we conclude that

$$
J^* - J^S \leq \sqrt{T}\bar{y}_{\infty} \sum_{l} r^l_{\max}\,,
$$

and the result follows.

$\square$

Algorithm 4 Static Probabilistic Control Heuristic (SPC)

```
Initialize c_1 ← C,
φ*_1 ← an optimal solution of Problem 3.2 with ρ = C/T
for t = 1 to T do    observe the opportunity class θ_t
  draw an action a_t with probability φ*_{ρt}(θ_t, a_t)
    if y(θ_t, a_t, ε) ≤ c_t, ∀ε ∈ ℰ, then choose the action a_t
      observe the shock ε_t
      c_{t+1} ← c_t − y(θ_t, a_t, ε_t)
    else choose action a_0.
```

C.1.2 Finite set of actions

Problem 3.8 is a linear program. In particular, introducing the set of slack variables $\{x_1 \dots x_L\}$, the standard form is given by

$$
\begin{aligned}
J(\rho) = \quad \max \quad & \sum_{\theta=1}^{\Theta} p_\theta \; \bar{r}_\theta \phi_\theta \\
\text{s.t} \quad & \sum_{\theta\in\Theta} p_\theta \; \bar{y}_{l\theta} \; \phi_\theta + x_l = \rho_l && \forall l \in [L] \\
& \sum_{a\in\mathcal{A}} \phi_\theta(a) = 1 && \forall \theta \in \Theta \\
& \phi_\theta(a) \geq 0 && \forall \theta \in \Theta, \forall a \in \mathcal{A} \\
& x_{\mathrm{i}} \geq 0 && \forall \mathrm{i} \in [L].
\end{aligned}
\tag{C.2}
$$

C.1.3 Dynamic bidding in repeated auctions

Second-price auctions

Lemma 12. *If f absolutely continuous and strictly increasing and f' is locally ξ-Lipschitz continuous in $\mathcal{N}(a_\theta^1, \varphi)$, then conditions CA 1- CA 6 hold.*

Proof. Let us see that conditions CA 1- CA 6 hold.

- **Conditions CA 1 and CA 2:** By hypothesis f is absolutely continuous and therefore both $\bar{q}$ and $\bar{m}$ are continuous.

- **Condition CA 3:** For each $\theta \in \Theta$, tet us define the function $G_\theta : \mathcal{A} \to \mathbb{R}$ by $G_\theta(a) = \theta f(a) - \int_0^a x \,\mathrm{d}f(x)$. Note that $G'_\theta(a) = (\theta - a)f'(a)$ and $\theta > 0$, then $\lim_{a\searrow 0} G'_\theta(a) > 0$ and $\beta(\theta) \neq 0$. If $\theta \neq \Theta_{\max}$, we also have $\lim_{a\to 0} G'_\theta(a) < 0$, and therefore $\beta(\theta) \in \arg\max_a G_\theta(a)$ is interior. Assume then first that $\theta \neq \Theta_{\max}$. In this case, we can compute the first order condition, obtaining that $\beta(\theta)$ satisfies the equation

  $$(\theta - \beta(\theta))f'(\beta(\theta)) = 0.$$

  Therefore, as the cumulative distribution function f is strictly increasing, the unique optimum is to bid truthfully, as it is known in the literature.

  Otherwise, if $\theta = \Theta_{\max}$, G_θ is strictly increasing for all $a \in \mathcal{A}$ and therefore $\beta(\theta) = \Theta_{\max}$.

  We then conclude that CA 3 holds.

- **Condition CA 4:** Due to the truthfulness property of the second price auction, from Proposition 9, it follows that $a_\theta^1 = \theta/(1+\mu^1)$ which belongs to $(0, \Theta_{\max})$ due to the bid is positive, and thus condition CA 4 holds.

- **Conditions CA 5 and CA 6:** Note first that $\bar{r}(\theta, a) = \theta f(a) - \bar{m}(\theta, a)$ and $\bar{y}(\theta, a) = \bar{m}(a)$, because f absolutely continuous. Then, it is enough to show the conditions hold for $h(\theta,\cdot)=\theta f(\cdot)$ and $\bar{m}(a)=\int_0^a x\,\mathrm{d}f(x)$. Specifically, we will show that if the density function f' is locally ξ-Lipschitz continuous in $\mathcal{N}(a^1_\theta,\varphi)$, then the gradient of $h(\theta,\cdot)=\theta f(\cdot)$ is locally $(\xi\Theta_{\max})$-Lipschitz continuous in $\mathcal{N}(a^1_\theta,\varphi)$ and $\bar{m}'(a)=af'(a)$ is locally $((\varphi+\Theta_{\max}/(\mu^1+1))\xi+\eta)-$Lipschitz continuous in $\mathcal{N}(a^1_\theta,\varphi)$.

To see the former note that

$$\|\nabla_a h(\theta,a)-\nabla_a h(\theta,a^1_\theta)\|=\theta|f'(a)-f'(a^1_\theta)|\leq\Theta_{\max}\xi|a-a^1_\theta|,$$

where the equality follows from the gradient of h and the inequality holds due to the locally $\xi-$Lipschitz continuity of f' and because $\theta\leq\Theta_{\max}$.

For the latter, we first show that $f'(a^1_\theta)\leq\eta$ with $\eta=1/\varphi+\xi\varphi$. Because the density f' is locally ξ-Lipschitz continuous in $\mathcal{N}(a^1_\theta,\varphi)$, we have that $f'(a^1_\theta)\leq f'(x)+\xi\varphi$ for all $x\in\in[a^1_\theta,a^1_\theta+\varphi]$. Integrating over $x\in[a^1_\theta,a^1_\theta+\varphi]$ we obtain that $f'(a^1_\theta)\varphi\leq 1+\xi\varphi^2$ because f' integrates to at most one. The result follows by dividing by φ. We now show that $\bar{m}'(a)$ is locally Lipschitz continuous:

$$\begin{aligned}|\bar{m}'(a)-\bar{m}'(a^1_\theta)| &= |af(a)-a^1_\theta f(a^1_\theta)|\\ &= |a\left[f'(a)-f'(a^1_\theta)\right]+f'(a^1_\theta)(a-a^1_\theta)|\\ &\leq a|f'(a)-f'(a^1_\theta)|+f'(a^1_\theta)|a-a^1_\theta|\\ &\leq \left(\left(\varphi+\frac{\Theta_{\max}}{\mu^1+1}\right)\xi+\eta\right)|a-a^1_\theta|,\end{aligned}$$

where the first inequality holds applying triangle inequality and the last follows from the bound of $f'(a^1_\theta)$, together with the locally $\xi-$Lipschitz continuity of f', the equality $a^1_\theta=\theta/(1+\mu^1)$ and the bound $a^1_\theta\leq\Theta_{\max}$. The proof is completed. □

First-price auctions

Lemma 13. *If f absolutely continuous, $M(a)=a+f(a)/f'(a)$ strictly increasing, the bid a^1_θ maximizing $\theta\bar{q}(a)-(\mu^1+1)\bar{m}(a)$ is interior, and the density function f' is locally $\xi-$Lipschitz in $\mathcal{N}(a^1_\theta,\varphi)$, and $f'(a^1_\theta)$ is upper bounded by η, conditions CA 1- CA 6 hold.*

Proof. As in the lemma for second-price auctions, conditions CA 1 and CA 2 holds because f is absolutely continuous. On the other hand, the bidder's problem in the static first price auction without budget constraints is to find a bid function $\beta(\theta)$ maximizing $(\theta-a)f(a)$. Note that $\arg\max\theta f(a)-af(a)=\arg\max\theta' f(a)-(\mu^1+1)af(a)$, where $\theta'=\theta/(\mu^1+1)$, is interior by hypothesis and then, computing the first order condition, we obtain that $\beta(\theta)$ should satisfy

$$f'(\beta(\theta))\theta-f(\beta(\theta))-\beta(\theta)f'(\beta(\theta))=0. \tag{C.3}$$

Then, we have $\theta=\beta(\theta)+f(\beta(\theta))/f'(\beta(\theta))=M(\beta(\theta))$ and by hypothesis we can compute the inverse of M and therefore $\beta(\theta)=M^{-1}(\theta)$. Thus, payments at the optimal solution are unique and assumption CA 3 holds.

Note that condition CA 4 is directly assumed in the statement of the lemma, and therefore it holds.

It remains to see smoothness of both the expected reward $\bar{r}(\theta, a) = (\theta - a)f(a)$ and expected payment $\bar{y}(a) = af(a)$ functions, but it is enough to show that the gradient of $\bar{y}(\theta, a)$ is locally Lipschitz continuous. To this end, note that

$$|\nabla_a \bar{y}(\theta, a)| = |f(a) + af(a) - f(a_\theta^1) - a_\theta^1 f(a_\theta^1)| \leq |f(a) - f(a_\theta^1)| + |af(a) - a_\theta^1 f'(a_\theta^1)|,$$

where the last expression in the inequality can be bound by using the local Lipschitz continuity of f' together with the upper bound for f' (as in the case of the second-price action, we have that $f'(a_\theta^1)$ is bounded) by using the mean value theorem. The remaining algebra is similar to the second-price case and the proof is completed. □

C.2 Proof Theorem 6

The goal of this section is to proof Theorem 6, which is our mean result regarding the performance of the heuristic CE for the set of DRCRC problems. To do that, we assume that Assumption 1 and Assumption 2 hold and we first introduce some processes and random variables, as well as technical results, that will be useful to obtain the desire result.

In what follows we will denote by y_t the resource consumption at time t if the decision maker follows the policy π^{CE}. That is, $y_t = y(\theta_t, a_t^{\pi^{\text{CE}}}, \varepsilon_t)$.

Let us consider the process $\{M_t\}_{t\geq 1}$ up to time T consisting in, at each time period, the accumulated difference between the resource vector consumption and its expectation, divided the remaining horizon. More specifically, for each $t \in [T]$,

$$M_t = \sum_{s=1}^{t} \frac{\mathbb{E}\left(y_s | \rho_s\right) - y_s}{T - s}.$$

Let us define the stopping time τ. To this end, we need to introduce two random variables. On one hand, we define τ_δ to be the first time t such that M_t has ℓ^2-norm greater or equal than δ, where δ is got form Assumption 2. That is,

$$\tau_\delta = \min_{t\in[T]} \{t : \|M_t\| \geq \delta\}.$$

If $\|M_t\|$ is at most δ for all $t \in [T]$, we set $\tau_\delta = \infty$.

On the other hand, we define τ_- as the first time at which there exists a resource such that its consumption under the policy $\phi^*_{\rho_t}$ is over capacity. That is,

$$\tau_- = \min_{t\in[T]} \{t : \exists\ l \in [L] \text{ s.t. } c_{t,l} - y_l(\theta_t, a^{\phi^*_{\rho_t}}, \varepsilon_t) < 0\}.$$

As above, if $c_{t,l} - y_l(\theta_t, a^{\phi^*_{\rho_t}}, \varepsilon_t)$ is greater or equal to 0 for all $t \in [T]$ and $l \in [L]$, we set $\tau_- = \infty$.

Then, we define the random variable τ as the minimum between τ_δ and τ_-, and the number of periods T, i.e., $\tau = \min\{\tau_\delta, \tau_-, T\}$.

Note that both τ_δ and τ_- are stopping times with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 1}$, with $\mathcal{F}_t = \sigma(\theta_1, \ldots, \theta_t, a_1, \ldots, a_t, \varepsilon_1, \ldots, \varepsilon_t)$, the history up to the end of period t, and thus we obtain that τ is also a stopping time with respect to the same filtration $\{\mathcal{F}_t\}_{t\geq 1}$.

Furthermore, the process $\{M_t\}_{t\geq 1}$ is a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 1}$. In fact, for each t, from Assumption 1.2 follows that $\mathbb{E}(y_t|\rho_t) - y_t \leq \bar{y}_\infty < \infty$, and therefore $\mathbb{E}(\|M_t\|) < \infty$ for all t. On the other hand, for each t holds that

$$M_{t+1} - M_t = \frac{\mathbb{E}(y_{t+1}|\rho_t) - y_{t+1}}{T - t - 1}$$

and

$$\mathbb{E}\left(\mathbb{E}(y_{t+1}|\rho_t) - y_{t+1}|\mathcal{F}_t\right) = 0,$$

concluding that

$$\mathbb{E}(M_{t+1} - M_t|\mathcal{F}_t) = 0 \ \forall t \geq 1.$$

Since $\{M_t\}_{t\geq 1}$ is a martingale and τ an stopping time, it turns out that the stopped process $\{M_{t\wedge\tau}\}_{t\geq 1}$ is also a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 1}$.

We are now ready to present some properties for the process and the stopping time defined above, that will be needed to prove the bound for the reward loss. The first of them states that up to time τ_δ, the random variable M_t can be expressed as the difference between ρ_{t+1} and ρ_1 and that ρ_t belongs to the ball centered at ρ_1 with radius δ.

Lemma 14. *Under Assumption 2, if t is at most τ_δ, it holds that:*

1. $\rho_{t+1} - \rho_1 = M_t$
2. $\|\rho_t - \rho_1\| < \delta$.

Proof. We will proceed by induction on t, dividing the proof into two steps, the first corresponds to prove the base case and the other the induction step.

Step 1. Note that for $t = 1$, the statement 2 of the lemma follows trivially and we are then under the hypothesis of Assumption 2, obtaining $\mathbb{E}(y_1|\rho_1) = \rho_1$. Therefore, we can express M_1 as follows:

$$M_1 = \frac{\mathbb{E}(y_1|\rho_1) - y_1}{T - 1} = \frac{\rho_1 - y_1}{T - 1}. \tag{C.4}$$

From the definition of ρ_1 and ρ_2, we have that $y_1 = \rho_1 T - \rho_2(T-1)$ and replacing in (C.4) follows that

$$M_1 = \frac{\rho_1 - \rho_1 T + \rho_2(T-1)}{T-1} = \rho_2 - \rho_1,$$

obtaining the first statement of the lemma and completes the step 1.

Step 2. Now, assume that Lemma 14 holds for all s smaller or equal than a fixed $t < \tau_\delta$ and let us prove that both statements also hold for $t+1$.

As in the base case, we will first prove the statement 2 and we then use it to prove statement 1. That is, let us show that $\|\rho_{t+1} - \rho_1\| < \delta$. Applying the induction hypothesis to t, it holds that $\|\rho_{t+1} - \rho_1\| = \|M_t\|$. On the other hand, $t < \tau_\delta$ and thus $\|\rho_{t+1} - \rho_1\| < \delta$, concluding that $\|M_t\| < \delta$, and the second statement follows.

In the remainder of the proof, we show that $\rho_{t+2} - \rho_1 = M_{t+1}$.

Note that

$$\rho_{t+2} - \rho_1 \quad = \quad \sum_{s=1}^{t+1} \rho_{s+1} - \rho_s = \sum_{s=1}^{t+1} \frac{\rho_s(T-s+1) - y_s}{T-s} - y_s = \sum_{s=1}^{t+1} \frac{\rho_s - y_s}{T-s} \ ,$$

where the first equality is obtained by using a telescoping sum and the second holds because $\rho_{s+1} = c_{s+1}/(T-s)$, $c_s = \rho_s(T-s+1)$ and $c_{s+1} = c_s - y_s$.

By the induction hypothesis, together with the statement 2 we already proved for $s = t+1$, it holds that $\|\rho_s - \rho_1\| < \delta$ for all $s \leq t+1$. Therefore we can apply Assumption 2 to the expression above obtaining that

$$\rho_{t+2} - \rho_1 = \sum_{s=1}^{t+1} \frac{\mathbb{E}(y_s|\rho_s) - y_s}{T-s} = M_{t+1},$$

and the lemma follows. $\square$

Since $\{M_{t\wedge\tau}\}_{t\geq 1}$ is a zero mean martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 1}$, a direct consequence of Lemma 14 is that also the stopped process $\{\rho_{t\wedge\tau}\}_{t\geq 1}$ is a martingale with respect to the same filtration.

The following lemma is a technical result we need to prove that the expected value of the remaining periods after the stopping time τ is bounded by a constant—result stated in Lemma 16. Specifically, it gives us sufficient conditions on t to be a lower bound for the stopping time τ_-.

Lemma 15. *Let assume that Assumption 1 and 2 hold, and define* $\Psi = \frac{\bar{y}_\infty}{\bar{\rho}_1 - \delta}$, *where* $\bar{\rho}_1$ *is the smallest component of vector* ρ_1, *and* $T^- = T + 1 - \Psi$. *If* $t \leq T^-$, *and* $t < \tau_\delta$ *then* $t < \tau_-$.

Proof. Due to the definition of the stopping time τ_-, we have to show that for all $s \leq t$, the consumption is at most the available capacity, i.e., $y(\theta_s, a^{\phi^*_{\rho_s}}, \varepsilon_s) \leq c_s$.

Take $s \leq t$. By hypothesis, $s < \tau_\delta$ and by Lemma 14 it holds that $\|\rho_s - \rho_1\| < \delta$. In particular, $|(\rho_s - \rho_1)_l| < \delta \; \forall l \in [L]$, obtaining

$$\rho_s > \rho_1 - \mathbf{1}\delta \,, \tag{C.5}$$

where $\mathbf{1}$ denotes the vector of ones of size L. On the other hand, note that

$$c_s > (T - s + 1)(\rho_1 - \mathbf{1}\delta) \geq \frac{\bar{y}_\infty}{\bar{\rho}_1 - \delta}(\rho_1 - \mathbf{1}\delta) \geq \mathbf{1}\bar{y}_\infty \geq y(\theta_s, a^{\phi^*_{\rho_s}}, \varepsilon_s),$$

where the strict inequality follows from the definition of ρ_s, together with inequality (C.5); the second inequality holds because $t \leq T^-$ and $\bar{\rho}_1 > \delta$, together with Assumption 1.2; and the third and the last due to the definition of $\bar{\rho}_1$ and $\bar{y}_\infty$, respectively.

We then conclude that τ_- is greater than t and the proof is completed.

□

Below we prove a result stating that the expected value of the remaining periods after the stopping time τ is upper bounded by a constant that does not depend on T, which is a key result to obtain the main theorem.

Lemma 16. *If Assumptions 1 and 2 hold, there exists a constant* μ *such that* $\mathbb{E}(T - \tau) \leq \mu$. *More specifically,*

$$\mathbb{E}(T - \tau) < \Psi + 14\frac{\bar{y}_\infty^2}{\delta^2}.$$

Proof. We will prove the result by bounding the expected value of τ, which is equivalent to the expression $\sum_{t=1}^{\infty} \mathbb{P}(\tau \geq t)$ because τ is a non-negative random variable. From the definition of τ, the probability of τ being greater than T is zero and the probability of being at least one is one, and then,

$$\mathbb{E}(\tau) = 1 + \sum_{t=2}^{T} \mathbb{P}(\tau_\delta \wedge \tau_- \geq t) = 1 + \sum_{t=2}^{T^- - 1} \mathbb{P}(\tau_\delta \wedge \tau_- \geq t) + \sum_{t=T^-}^{T} \mathbb{P}(\tau_\delta \wedge \tau_- \geq t) \geq 1 + \sum_{t=2}^{T^- - 1} \mathbb{P}(\tau_\delta \wedge \tau_- \geq t) \tag{C.6}$$

where the last equality follows just splitting the horizon and the inequality holds because $\mathbb{P}(\tau_\delta \wedge \tau_- \geq t) \geq 0$.

On the other hand,

$$\begin{aligned}
\sum_{t=2}^{T^- -1} \mathbb{P}(\tau_\delta \wedge \tau_- \geq t) &= \sum_{t=2}^{T^- -1} \mathbb{P}\left(\min_{s\in[T]} \{s : \|M_s\| \geq \delta\} \geq t \right) \\
&= \sum_{t=2}^{T^- -1} \mathbb{P}\left(\|M_s\| < \delta \ \forall s \in [t] \right) \\
&= T^- - 2 - \sum_{t=2}^{T^- -1} \mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta \right) \ ,
\end{aligned}$$

where the first equality is obtained by Lemma 15 (since $t < T^-$, $\tau_\delta \wedge \tau_- = \tau_\delta$) and the last one because $\mathbb{P}(\|M_s\| < \delta \ \forall s \in [t]) = 1 - \mathbb{P}(\max_{s\in[t]} \|M_s\| \geq \delta)$.

Then, using the equality above in (C.6) it holds that

$$\mathbb{E}(\tau) \geq T^- - 1 - \sum_{t=2}^{T^- -1} \mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta \right).$$

In the remainder of the proof we will upper bound $\sum\limits_{t=2}^{T^- -1} \mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta \right)$, and we proceed by applying Theorem 3.5 in [115]. To this end, note first that $(\mathbb{R}^L, \|\cdot\|)$ is a separable Banach space, and since $\|x+y\| + \|x-y\| \leq 2\|x\|^2 + 2\|y\|^2$ holds for all $x, y \in \mathbb{R}_+$, it is $(2,1)$-smooth. Define, for each t, the Martingale $\{M_{t\wedge s}\}_{s\geq 1}$. From the definition of M_s follows that

$$M_s - M_{s-1} = \frac{\mathbb{E}(y_s|\rho_s) - y_s}{T-s},$$

and therefore

$$\begin{aligned}
\sum_{s=1}^{\infty} \|M_s - M_{s-1}\|_\infty^2 &= \sum_{s=1}^{t} \|M_s - M_{s-1}\|_\infty^2 \\
&= \sum_{s=1}^{t} \left\| \frac{\mathbb{E}(y_s|\rho_s) - y_s}{T-s} \right\|_\infty^2 \\
&\leq (2\bar{y}_\infty)^2 \frac{1}{T-t},
\end{aligned}$$

where the inequality follows from Assumption 1.2, and using that $\sum_{s=1}^{t} 1/(T-s)^2 \leq \int_0^t 1/(T-s)^2 < 1/(T-t)$.

Then, we are under the hypothesis of the theorem mentioned above, and applying it together with the inequality $\mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta \right) \leq 1$, we obtain

$$\mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta \right) \leq 1 \wedge 2\exp\left(-\frac{\delta^2 (T-t)}{8\bar{y}_\infty^2} \right).$$

Summing on t and using the bound obtained above, we have

$$\begin{aligned}\sum_{t=2}^{T^- -1} \mathbb{P}\left(\max_{s\in[t]} \|M_s\| \geq \delta\right) &\leq \sum_{t=2}^{T}\left(2\exp\left(-\frac{\delta^2(T-t)}{8\bar{y}_\infty^2}\right)\wedge 1\right)\\ &\leq \int_0^T \left(2\exp\left(-\frac{\delta^2(T-t)}{8\bar{y}_\infty^2}\right)\wedge 1\right)\mathrm{d}t\\ &\leq \frac{8\bar{y}_\infty^2}{\delta^2}(\log 2+1),\end{aligned}$$

where the second inequality follows from bounding the summation by the integration and the last inequality from Lemma 17.

Putting all together and bounding $8(\log 2 + 1)$ by 14 we conclude

$$\begin{aligned}\mathbb{E}(T-\tau) &< T - T^- + 1 + 14\frac{\bar{y}_\infty^2}{\delta^2}\\ &= \Psi + 14\frac{\bar{y}_\infty^2}{\delta^2},\end{aligned}$$

and the desire result is obtained. □

We are now ready to prove Theorem 6 by combining the technical results already presented.

Proof of Theorem 6. We have to bound $J^* - J^{\mathrm{CE}}$, which is upper bounded by $TJ(\rho_1) - J^{\mathrm{CE}}$ because $J^* \leq TJ(C/T)$ (see, e.g., [66]). Thus, it is enough to bound $TJ(\rho_1) - J^{\mathrm{CE}}$.

By dividing the horizon from 1 to τ and from τ to T, we obtain

$$TJ(\rho_1) - J^{\mathrm{CE}} \leq \underbrace{\mathbb{E}\left(\sum_{t=1}^{\tau} J(\rho_1) - \sum_{t=1}^{\tau} r(\theta_t, a_t^{CE}, \varepsilon_t)\right)}_{(A)} + \underbrace{\mathbb{E}\left(\sum_{t=\tau+1}^{T} J(\rho_1)\right)}_{(B)}, \tag{C.7}$$

and we then have to bound (A) and (B), which will be done in Part 1 and Part 2 separately, respectively.

Part 1. We will prove that $(A) \leq K/2\bar{y}_2^2 \log T$, and we divide the proof into three steps. First, we show that the expected reward earned up to time τ considering the policy given by the CE heuristic equals the expected reward until time τ of the deterministic problem for

$\rho = \rho_t$ at time t. Using that, applying Assumption 2- the hypothesis is fulfilled because $t \leq \tau$ and then by Lemma 14 we have $\|\rho_t - \rho_1\| < \delta$-, summing over t and taking expectation, we obtain

$$(A) = \mathbb{E}\left(\sum_{t=1}^{\tau} (J(\rho_1) - J(\rho_t))\right) \leq \underbrace{\mathbb{E}\left(\sum_{t=1}^{\tau} -\nabla J(\rho_1)(\rho_t - \rho_1)\right)}_{(A_1)} + \underbrace{\mathbb{E}\left(\sum_{t=1}^{\tau} \frac{K}{2}\|\rho_t - \rho_1\|^2\right)}_{(A_2)}.$$

In the second step we bound (A_1) and in step 3 we bound (A_2).

Step 1. Let us prove that

$$\mathbb{E}\left(\sum_{t=1}^{\tau} r(\theta_t, a_t^{CE}, \varepsilon_t)\right) = \mathbb{E}\left(\sum_{t=1}^{\tau} J(\rho_t)\right). \tag{C.8}$$

To this end, consider the sequence of zero mean, i.i.d. random variables $\{X_t\}_{t\geq 1}$ given by

$$X_t = r(\theta_t, a_t^{CE}, \varepsilon_t) - \mathbb{E}_{\theta,\varepsilon}\left(r(\theta_t, a_t^{CE}, \varepsilon_t)|\rho_t\right).$$

Then, it is well known (see e.g. [118] page 296) that defining $N_s = \sum_{t=1}^{s} X_t$, holds that $\{N_s\}_{s\geq 1}$ is a martingale relative to the filtration $\{\mathcal{F}_t\}_{t\geq 1}$ previously defined. Therefore, due to τ is an stopping time with respect to the same filtration, we can apply the Martingale Stopping Theorem ([118], Theorem 6.6.2), which in turns implies that

$$\mathbb{E}\left(N_\tau\right) = \mathbb{E}\left(N_1\right) = 0. \tag{C.9}$$

On the other hand, by the definition of the deterministic problem, we have that

$$\mathbb{E}_{\theta,\varepsilon}\left(r(\theta_t, a^{CE}, \varepsilon_t)|\rho_t\right) = J(\rho_t) \quad \forall t \in [T]. \tag{C.10}$$

Then, (C.8) follows easily from (C.9) and (C.10).

Step 2. We want to bound (A_1). In fact, we will prove that it is equal to zero. From the linearity of the expectation, (A_1) is equivalent to

$$\sum_{t=1}^{\tau} -\nabla J(\rho_1)\mathbb{E}\left(\rho_t - \rho_1\right).$$

Furthermore, note that $\mathbb{E}\left(\rho_t - \rho_1\right)$ is zero because $\{\rho_{t\wedge\tau}\}_{t\geq 1}$ is a martingale, and the term (A_1) vanishes in the bound, obtaining $(A) \leq (A_2)$.

Step 3. Regarding (A_2), by linearity of the expectation, it is enough to bound the expression $\mathbb{E}\left(||\rho_t - \rho_1||^2\right)$, for $t \leq \tau$. Using telescoping sum and the orthogonality of the martingale's increments (see e.g. [78], Chapter 10 Lemma 4.1) we have

$$\mathbb{E}\left(\|\rho_t - \rho_1\|^2\right) = \mathbb{E}\left\|\sum_{s=2}^{t}(\rho_s - \rho_{s-1})\right\|^2 = \sum_{s=2}^{t}\mathbb{E}\|\rho_s - \rho_{s-1}\|^2, \tag{C.11}$$

and therefore in the remainder of this step we bound $\|\rho_s - \rho_{s-1}\|^2$. Note that

$$\begin{aligned}
\|\rho_s - \rho_{s-1}\|^2 &= \left\|\frac{\rho_{s-1}(T-s+2) - y_{s-1}}{T-s+1} - \rho_{s-1}\right\|^2 \\
&= \left\|\frac{\rho_{s-1} - y_{s-1}}{T-s+1}\right\|^2 \\
&= \left\|\frac{\mathbb{E}(y_{s-1}|\rho_{s-1}) - y_{s-1}}{T-s+1}\right\|^2,
\end{aligned}$$

where the first equality follows from the definitions $\rho_s = \frac{c_s}{T-s+1}$, $\rho_{s-1} = \frac{c_{s-1}}{T-s+2}$ and $c_s = c_{s-1} - y_{s-1}$; and the last equality follows from Assumption 2.

Furthermore, by definition of ℓ^2-norm, holds that

$$\|\mathbb{E}(y_{s-1}|\rho_{s-1}]) - y_{s-1}\|^2 = \sum_{l\in[L]}\left(\mathbb{E}(y_{s-1,l}|\rho_{s-1}) - y_{s-1,l}\right)^2,$$

and taking expectation follows that

$$\mathbb{E}\left(\|\mathbb{E}(y_{s-1}|\rho_{s-1}) - y_{s-1}\|^2\right) = \sum_{l\in[L]}\mathbb{E}\left((\mathbb{E}(y_{s-1,l}|\rho_{s-1}) - y_{s-1,l})^2\right) = \sum_{l\in[L]}\mathrm{Var}(y_{s-1,l}).$$

Using that $\mathrm{Var}(y_{s-1,l}) = \mathbb{E}\left((y_{s-1,l})^2\right) - \mathbb{E}\left(y_{s-1,l}\right)^2 \leq \mathbb{E}\left((y_{s-1,l})^2\right)$, together with the linearity of the expectation we have

$$\sum_{l\in[L]}\mathrm{Var}(y_{s-1,l}) \leq \sum_{l\in[L]}\mathbb{E}\left((y_{s-1,l})^2\right) = \mathbb{E}\left(\sum_{l\in[L]}(y_{s-1,l})^2\right) \leq \bar{y}_2^2,$$

where the last inequality follows from Assumption 1.2.

Getting back to expression (C.11), we finally obtain

$$\mathbb{E}\left(\|\rho_t - \rho_1\|^2\right) \leq \sum_{s=2}^{t}\frac{\bar{y}_2^2}{(T-s+1)^2} \leq \bar{y}_2^2\int_2^t \frac{1}{(T-s+1)^2}\,\mathrm{d}s \leq \frac{\bar{y}_2^2}{T-t+1},$$

and therefore by doing the suitable computations we obtain $(A_2) \leq K/2\bar{y}_2^2 \log T$, concluding the proof of Step 3.

Putting all together we get

$$\mathbb{E}\left(\sum_{t=1}^{\tau}(J(\rho_1) - J(\rho_t))\right) \leq \frac{K}{2}\bar{y}_2^2 \log T\,, \tag{C.12}$$

and the proof of Part 1 is complete.

Part 2. It only remains to bound the second term in (C.7). Note that

$$\mathbb{E}\left(\sum_{t=\tau+1}^{T} J(\rho_1)\right) \quad = \quad \mathbb{E}(T-\tau)J(C/T),$$

and applying Lemma 16 we obtain

$$\mathbb{E}\left(\sum_{t=\tau+1}^{T} J(\rho_1)\right) \leq \left[\Psi + 14\frac{\bar{y}_\infty^2}{\delta^2}\right] J(C/T). \tag{C.13}$$

Using (C.12) together with (C.13) in (C.7) we get

$$J^* - J^{CE} \leq \bar{y}_2^2 K \log T + \left[\Psi + 14\frac{\bar{y}_\infty^2}{\delta^2}\right] J(C/T),$$

and the result follows.

□

C.3 Proofs for Section 3.7

C.3.1 Proof of Proposition 7

We divide the proof into two steps. First, we prove that an optimal solution to the dual problem exists, namely μ^*. Then, we define ϕ^* properly and we apply Proposition 5.1.5 in [23] to prove that ϕ^* is primal solution and μ^* is in fact a Lagrangian multiplier and that therefore there is no duality gap, obtaining the desire result.

Step 1. Note that for each $\theta \in \Theta$, g_θ is convex because it is defined as the supremum of a family of linear functions and therefore the dual problem is a convex problem.

To prove the existence of optimal dual solution μ^*, we first prove Ψ_ρ is differentiable (and thus continuous) and then we argue that the domain of the dual problem can be restricted to a compact set, achieving the result applying the extreme value theorem. For the former, it is enough to note that $g_\theta(\mu)$ is differentiable by Assumption SC 1, and in particular we have that Ψ_ρ is continuous. On the other hand, we can prove that we can restrict the domain of the dual problem to the hypercube $[0, \bar{\mu}_\rho]^L$, for $\bar{\mu}_\rho = \bar{r}_\infty/\bar{\rho}$, where $\bar{\rho} = \min_{l\in[L]} \rho_l$ and $\bar{r}_\infty$ is the positive real number provided by Assumption1.1. We have that $\bar{\mu}_\rho < \infty$ because $\rho > 0$. Let us check that every $\mu \notin [0, \bar{\mu}_\rho]^L$ is suboptimal. Take $\mu \notin [0, \bar{\mu}_\rho]^L$, and define $L_1 = \{l \in [L] : \mu_l > \bar{\mu}_\rho\}$ the components of μ greater than $\bar{\mu}_\rho$. Then, we have

$$\Psi_\rho(\mu) \geq \rho^\top \mu \geq \sum_{l\in L_1} \rho_l \bar{\mu}_l = \sum_{l\in L_1} \bar{r}_\infty \frac{\rho_l}{\bar{\rho}} \geq \bar{r}_\infty \geq \Psi_\rho(0),$$

where the first inequality holds because $\bar{r}(\theta, a_0) = \bar{y}(\theta, a_0) = 0$, for all $\theta \in \Theta$ and therefore $g_\theta(\mu) \geq 0$; the second follows from the non-negativity of vectors μ and ρ, the third inequality holds because L_1 contains at least one element and $\rho_l \geq \bar{\rho}$, and the last one follows because $\Psi_\rho(0) = \sum_{\theta\in\Theta} p_\theta \max_{a\in\mathcal{A}} \bar{r}(\theta, a)$ and $\bar{r}_\infty \geq \bar{r}(\theta, a)$ for all $\theta \in \Theta$, and $a \in \mathcal{A}$. Then, we have $\Psi_\rho(0) \leq \Psi_\rho(\mu)$ and together with the extreme value theorem we conclude that for each $\rho > 0$ there exist μ^* optimal dual solution satisfying $\mu^* \in [0, \bar{\mu}_\rho]^L$.

Step 2. Given $\rho > 0$, take μ^* an optimal dual solution and, for each $\theta \in \Theta$, define ϕ^*_θ a distribution that assigns probability one to an action $a^*_\theta \in \arg\max_{a\in\mathcal{A}} \left\{\bar{r}(\theta, a) - {\mu^*}^\top \bar{y}(\theta, a)\right\}$. Such actions are guaranteed to exist by Assumption SC 1. Let us now show that (ϕ^*, μ^*) is an optimal solution-Lagrange multiplier pair. We will proceed by using Proposition 5.1.5 in [23]. That is, we need to check primal and dual feasibility, Lagrangian optimality and complementary slackness.

1. *Primal and dual feasibility.* Dual feasibility follows because $\mu^* \geq 0$. For primal feasibility, note that from the envelope theorem applied to g_θ (see, e.g., Theorem 1 in [108]), the gradient of Ψ_ρ evaluated at μ^* is given by

$$\nabla\Psi_\rho(\mu^*) = \rho + \sum_{\theta\in\Theta} p_\theta \nabla g_\theta(\mu^*) = \rho - \sum_{\theta\in\Theta} p_\theta \bar{y}(\theta, a^*_\theta)\,, \tag{C.14}$$

   where we used that by SC 1 the value function $g_\theta(\mu)$ is differentiable and achieved for an action a^*_θ, and that the gradient of $\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)$ with respect to μ exists and is given by $\bar{y}(\theta, a)$. Because μ^* is an optimal dual solution and the constraint set is convex, by Proposition 2.1.2 in [23], the first-order conditions are given by

$$\nabla\Psi_\rho(\mu^*)^\top (\mu - \mu^*) \geq 0,\ \forall \mu \in \mathbb{R}^L_+\,. \tag{C.15}$$

   Letting $\mu_l \to \infty$, we obtain that $\nabla\Psi_\rho(\mu^*) \geq 0$, which, in turn, implies that $\sum_{\theta\in\Theta} p_\theta \bar{y}(\theta, a^*_\theta) \leq \rho$ by (C.14). Primal feasibility follows.

2. *Complementary slackness.* If $\mu^*_l = 0$, we trivially have $(\nabla\Psi_\rho(\mu^*))_l \mu^*_l = 0$ and complementary slackness follows. If $\mu^*_l > 0$, we can take $\nu > 0$ with $\mu^*_l + \nu$ and $\mu^*_l - \nu$

belonging to $\mathbb{R}_+$. Using (C.15), we obtain that $(\nabla\Psi_\rho(\mu^*))_l\nu \geq 0$ and $(\nabla\Psi_\rho(\mu^*))_l\nu \leq 0$. Thus, it holds that $(\nabla\Psi_\rho(\mu^*))_l = 0$ and complementary slackness follows.

3. *Lagrangian optimality.* Note that

$$\begin{aligned}
\arg\max_{\phi\in\Phi} \mathcal{L}(\phi,\mu^*) &= \arg\max_{\phi\in\Phi} \left\{ \mu^*\rho + \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} \left(\bar{r}(\theta,a) - {\mu^*}^\top \bar{y}(\theta,a) \right) \mathrm{d}\phi_\theta(a) \right\} \\
&= \left\{ \arg\max_{\phi_\theta\in\Delta(\mathcal{A})} \int_{\mathcal{A}} \left(\bar{r}(\theta,a) - {\mu^*}^\top \bar{y}(\theta,a) \right) \mathrm{d}\phi_\theta(a) \right\}_{\theta\in\Theta} \\
&= \left\{ \arg\max_{a\in\mathcal{A}} g_\theta(\mu^*) \right\}_{\theta\in\Theta},
\end{aligned}$$

where the second equality holds because we can separate the problem for each θ. But note that $g_\theta(\mu^*)$ is maximized at a^*_θ and thus we have Lagrangian optimality.

Therefore, the four conditions hold and the proof is complete.

□

C.3.2 Proof of Lemma 7

We have to show that if $\rho \in \mathcal{N}(\rho_1, \delta)$, with $\delta = (\nu\kappa)/2$ then it holds that

1. $J(\rho) \geq J(\rho_1) + \nabla J(\rho_1)(\rho - \rho_1) - \frac{1}{2\kappa}\|\rho - \rho_1\|^2$,

2. $\sum_{\theta\in\Theta} p_\theta \int_a \bar{y}(\theta,a)\,\mathrm{d}\phi^*_\theta(a) = \rho$.

Part 1. We first extend the strong convexity lower bound of g_θ to the entire domain. Given $\theta \in \Theta$, by SC 2, g_θ admits a κ-LUQ envelope in $I^\nu = \mathcal{N}(\mu^1, \nu)$. Then, for all $\mu \in I^\nu$.

$$g_\theta(\mu) \geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \frac{\kappa}{2}\|\mu - \mu^1\|^2 \,. \tag{C.16}$$

We next extend the lower bound to every feasible dual variable. Consider $\mu \geq 0$ with $\mu \notin I^\nu$. Take α such that $\alpha\mu + (1-\alpha)\mu^1 = \hat{\mu}$ where $\hat{\mu}$ is in the boundary of the ball I^ν, i.e., $\hat{\mu} \in \partial I^\nu$. The latter is possible because $\mu \notin I^\nu$. Note that $\hat{\mu} - \mu^1 = \alpha(\mu - \mu^1)$. Taking ℓ^2-norm in both sides, we get that $\alpha = \|\hat{\mu} - \mu^1\|/\|\mu - \mu^1\|$. Moreover, $\alpha \in (0,1)$ because μ^1 is interior since $\nu > 0$ and $\mu \notin I^\nu$. Because g_θ is convex, we have

$$\alpha g_\theta(\mu) + (1-\alpha) g_\theta(\mu^1) \geq g_\theta(\alpha\mu + (1-\alpha)\mu^1) = g_\theta(\hat{\mu}) \,,$$

which can be reordered to give

$$
\begin{aligned}
g_\theta(\mu) &\geq \frac{1}{\alpha} g_\theta(\hat{\mu}) - \frac{1-\alpha}{\alpha} g_\theta(\mu^1) \\
&\geq \frac{1}{\alpha} g_\theta(\mu^1) + \frac{1}{\alpha}\nabla g_\theta(\mu^1)^\top(\hat{\mu} - \mu^1) + \frac{\kappa}{2\alpha}\|\hat{\mu} - \mu^1\|^2 - \frac{1-\alpha}{\alpha} g_\theta(\mu^1) \\
&= g_\theta(\mu^1) + \frac{1}{\alpha}\nabla g_\theta(\mu^1)^\top(\hat{\mu} - \mu^1) + \frac{\kappa}{2\alpha}\|\hat{\mu} - \mu^1\|^2 \,, \\
&= g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \frac{\kappa\nu}{2}\|\mu - \mu^1\| \,,
\end{aligned}
$$

where the second inequality follows by (C.16) with $\mu = \hat{\mu}$ and the second equality from $\hat{\mu} - \mu^1 = \alpha(\mu - \mu^1)$ together with $\|\hat{\mu} - \mu^1\|^2 = \alpha\|\mu - \mu^1\|\|\hat{\mu} - \mu^1\| = \alpha\nu\|\mu - \mu^1\|$ since $\|\hat{\mu} - \mu^1\| = \nu$ because $\hat{\mu}$ lies at the boundary of the ball I^ν. Combining both cases we obtain that

$$g_\theta(\mu) \geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \kappa\ell(\mu - \mu^1) \,, \tag{C.17}$$

where

$$\ell(z) = \begin{cases} \frac{1}{2}\|z\|^2 & \text{if } \|z\| \leq \nu \\ \frac{\nu}{2}\|z\| & \text{otherwise.} \end{cases}$$

The function $\ell : \mathbb{R}^L \to \mathbb{R}$ is, unfortunately, not convex. We restore convexity while preserving the lower bound by shrinking the radius of ball in half and shifting down the cone outside the ball. In particular, consider the function $f^* : \mathbb{R}^L \to \mathbb{R}$ given by

$$f^*(z) = \begin{cases} \frac{1}{2}\|z\|^2 & \text{if } \|z\| \leq \frac{\nu}{2} \\ \frac{\nu}{2}\|z\| - \frac{1}{8}\nu^2 & \text{otherwise.} \end{cases}$$

The function is easily shown to be convex and satisfies $\ell(z) \geq f^*(z)$ for all $z \in \mathbb{R}^L$. (Actually, $f^*(z)$ is the largest convex function satisfying $\ell(z) \geq f^*(z)$.) Putting everything together we obtain that

$$g_\theta(\mu) \geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \kappa f^*(\mu - \mu^1) \,. \tag{C.18}$$

Part 2. Let us prove the first statement. Using this lower bound on g_θ to bound $J(\rho)$ we have

$$
\begin{aligned}
J(\rho) &\geq \min_{\mu\geq 0}\left\{\rho^\top\mu + \sum_{\theta\in\Theta} p_\theta\left(g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \kappa f^*(\mu - \mu^1)\right)\right\} \\
&= \rho_1^\top\mu^1 + \sum_{\theta\in\Theta} p_\theta g_\theta(\mu^1) + (\mu^1)^\top(\rho - \rho_1) + \min_{\mu\geq 0}\left\{(\rho - \rho_1)^\top(\mu - \mu^1) + \kappa f^*(\mu - \mu^1)\right\} \\
&= J(\rho_1) + \nabla J(\rho_1)^\top(\rho - \rho_1) + \underbrace{\min_{z\geq\mu^1}\left\{(\rho - \rho_1)^\top z + \kappa f^*(z)\right\}}_{(E)} \,,
\end{aligned}
$$

where the first equality holds because the dual solution is interior, i.e., $\mu^1 > 0$, and then the first order conditions for the dual problem imply that $\rho_1 + \sum_{\theta\in\Theta} p_\theta \nabla g_\theta(\mu^1) = 0$ and the last equality follows from performing the change of variables $\mu - \mu^1 = z$ and because $J(\rho_1) = \rho_1^\top \mu^1 + \sum_{\theta\in\Theta} p_\theta g_\theta(\mu^1)$ together with $\nabla J(\rho_1) = \mu^1$ from the envelope theorem. Note that envelope theorem applies to J because both g_θ and J–in a neighborhood of ρ_1–are continuously differentiable (it follows from Assumption SC 1, and from the concavity of J and Theorem 25.5 in [117], respectively).

In the remainder of the proof we lower bound the error term (E). We have that

$$(E) \geq \min_{z\in\mathbb{R}^L} \left\{(\rho-\rho_1)^\top z + \kappa f^*(z)\right\} = -\kappa \max_{z\in\mathbb{R}^L} \left\{ \left(\frac{\rho_1-\rho}{\kappa}\right)^\top z - f^*(z) \right\} = -\kappa f^{**}\left(\frac{\rho_1-\rho}{\kappa}\right),$$

where the first inequality follows from relaxing the constraint that $z \geq -\mu^1$, the first equality from factoring $\kappa > 0$ and changing the direction of the optimization, and the last one by denoting $f^{**}(x) = \max_{z\in\mathbb{R}^L}\{x^\top z - f^*(z)\}$ to be the convex conjugate of $f^*(z)$. Invoking Lemma 18 with $\varphi = \nu/2$, we obtain that $f^{**}(x) = f(x)$ with $f(x) = \frac{1}{2}\|x\|^2$ if $\|x\| \leq \nu/2$ and $f(x) = \infty$ otherwise because the function $f(x)$ is proper (because $\nu > 0$), closed, and convex (because every squared norm is convex). Therefore, if $\|\rho - \rho_1\| \leq \nu\kappa/2$, we have

$$(E) \geq -\kappa f\left(\frac{\rho_1-\rho}{\kappa}\right) = -\frac{1}{2\kappa}\|\rho-\rho_1\|^2.$$

Putting it all together, we conclude that for ρ such that $\|\rho - \rho_1\| \leq (\nu\kappa)/2$,

$$J(\rho) \geq J(\rho_1) + \nabla J(\rho_1)^\top(\rho-\rho_1) - \frac{1}{2\kappa}\|\rho-\rho_1\|^2.$$

Part 3. Let us now see the second statement. More specifically, we will show that if $||\rho_1 - \rho|| \leq \delta$ then

$$\sum_{\theta\in\Theta} p_\theta \int_{a\in\mathcal{A}} \bar{y}(\theta,a)\,\mathrm{d}\phi_\theta^*(a) = \rho,$$

where ϕ^* is an optimal solution of the deterministic problem when the resource vector is ρ.

Note that by complementary slackness we know that for all i $\in [L]$

$$\mu_{\mathrm{i}}\left(\rho - \sum_{\theta\in\Theta} p_\theta \int_{a\in\mathcal{A}} \bar{y}(\theta,a)\,\mathrm{d}\phi_\theta^*(a)\right)_{\mathrm{i}} = 0,$$

where μ is the optimal solution of (3.5). That is,

$$\mu \in \arg\min_{\mu\in\mathbb{R}_+^L} \Psi_\rho(\mu),$$

with $\Psi_\rho(\mu) = \rho^\top \mu + \sum_{\theta\in\Theta} p_\theta g_\theta(\mu)$ the Lagrange dual function. We prove the result by showing that $\mu_{\mathrm{i}} > 0$ $\forall \mathrm{i} \in [L]$ for every optimal solution when the resource vector ρ satisfies $\|\rho - \rho_1\| < \delta$.

For all $\mu \geq 0$, we have from (C.17) that

$$\begin{aligned}
\Psi_\rho(\mu) &\geq \rho^\top \mu + \sum_{\theta\in\Theta} p_\theta \left(g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top (\mu - \mu^1) + \kappa \ell(\mu - \mu^1) \right) \\
&= \rho^\top \mu + \sum_{\theta\in\Theta} p_\theta g_\theta(\mu^1) + (\mu^1 - \mu)^\top \rho_1 + \kappa \ell(\mu - \mu^1) \\
&= \Psi_\rho(\mu^1) + \kappa \ell(\mu - \mu^1) + (\mu^1 - \mu)^\top (\rho_1 - \rho).
\end{aligned}$$

Let us define $U_0 = \left\{ \mu \in \mathbb{R}_+^L : \mu_j = 0 \text{ for some } j \in [L] \right\}$ the set of dual feasible solutions with some zero component. We will prove that if ρ is suitably chosen, then $\min_{\mu\in U_0} \Psi_\rho(\mu) > \Psi_\rho(\mu^1)$ and therefore the optimal solution of the dual problem satisfies $\mu_{\mathrm{i}} > 0$ for all $\mathrm{i} \in [L]$ since all dual solutions in the boundary U_0 are strictly dominated by μ^1. To this end, it is sufficient to show that

$$(I) = \min_{\mu\in U_0} \left\{ \kappa \ell(\mu - \mu^1) + (\mu^1 - \mu)^\top (\rho_1 - \rho) \right\} > 0 \,.$$

Suppose that $\|\mu - \mu^1\| \leq \nu$. In this case, $\ell(z) = \frac{1}{2}\|z\|^2$. Before proceeding we note that $\min_{\mu\in U_0} \|\mu - \mu^1\| \geq \underline{\mu}$. In fact, if $\mu \in U_0$ there exist $j \in [L]$ such that $\mu_j = 0$ and therefore $(\mu - \mu^1)_j = -\mu^1_j$, obtaining that $\|\mu - \mu^1\| \geq \mu^1_j \geq \underline{\mu}$. Then, using Cauchy-Schwartz we obtain that

$$\begin{aligned}
(I) &\geq \min_{\mu\in U_0} \left\{ \frac{\kappa}{2} \|\mu - \mu^1\|^2 - \|\rho_1 - \rho\| \|\mu - \mu^1\| \right\} = \min_{\mu\in U_0} \|\mu - \mu^1\| \left(\frac{\kappa}{2} \|\mu^1 - \mu\| - \|\rho_1 - \rho\| \right) \\
&\geq \min_{\mu\in U_0} \|\mu - \mu^1\| \left(\frac{\kappa \underline{\mu}}{2} - \|\rho_1 - \rho\| \right) \geq \underline{\mu} \left(\frac{\kappa \underline{\mu}}{2} - \|\rho_1 - \rho\| \right) > 0 \,,
\end{aligned}$$

where the second inequality follows because $\min_{\mu\in U_0} \|\mu - \mu^1\| \geq \underline{\mu}$, the third and fourth inequalities because $\|\rho_1 - \rho\| < (\nu\kappa)/2 \leq (\underline{\mu}\kappa)/2$ because $\nu \leq \underline{\mu}$.

Suppose that $\|\mu - \mu^1\| > \nu$. In this case, $\ell(z) = \nu\|z\|/2$ and using again Cauchy-Schwartz we obtain

$$(I) \geq \min_{\mu\in U_0} \left(\frac{\kappa\nu}{2} - \|\rho_1 - \rho\| \right) \|\mu - \mu^1\| \geq \left(\frac{\kappa\nu}{2} - \|\rho_1 - \rho\| \right) \nu > 0 \,,$$

where the third inequality follows because $\|\rho - \rho_1\| < (\kappa\nu)/2$. The result follows.

$\square$

C.3.3 Proof of Lemma 8

We will show that under assumptions CA 1-CA 6, $g_\theta(\mu)$ admits a κ-LUQ envelope in $\mathcal{N}(\mu^1, \nu)$ for $\nu = \kappa\varphi/\sigma$ and $\kappa = \kappa_r + (\nu + \|\mu^1\|)\|\kappa_y\|$.

From CA 5, $\bar{r}(\theta, \cdot)$ admits a κ_r-LDQ envelope in $\mathcal{N}(a^1_\theta, \varphi)$. That is, for all $\theta \in \Theta$,

$$\bar{r}(\theta, a) \geq \bar{r}(\theta, a^1_\theta) + \nabla\bar{r}(\theta, a^1_\theta)^\top (a - a^1_\theta) - \frac{\kappa_r}{2}\|a - a^1_\theta\|^2 \quad \forall a \in \mathcal{N}(a^1_\theta, \varphi).$$

On the other hand, from CA 6, $\bar{y}(\theta, \cdot)$ admits a κ_y-UUQ in $\mathcal{N}(a^1_\theta, \varphi)$. That is, for all $\theta \in \Theta$,

$$\bar{y}(\theta, a) \leq \bar{y}(\theta, a^1_\theta) + \nabla\bar{y}(\theta, a^1_\theta)(a - a^1_\theta) + \frac{\kappa_y}{2}\|a - a^1_\theta\|^2 \quad \forall a \in \mathcal{N}(a^1_\theta, \varphi).$$

Combining these two inequalities we obtain that, for $a \in \mathcal{N}(a^1_\theta, \varphi)$, we have

$$\begin{aligned}
&\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a) \\
&\quad \geq \bar{r}(\theta, a^1_\theta) - \mu^\top \bar{y}(\theta, a^1_\theta) + (\nabla\bar{r}(\theta, a^1_\theta) - \nabla\bar{y}(\theta, a^1_\theta)^\top \mu)^\top (a - a^1_\theta) - \frac{\kappa_r + \mu^\top \kappa_y}{2}\|a - a^1_\theta\|^2 \\
&\quad \geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top (\mu - \mu^1) + (\nabla\bar{y}(\theta, a^1_\theta)^\top (\mu^1 - \mu))^\top (a - a^1_\theta) - \frac{\kappa_r + \mu^\top \kappa_y}{2}\|a - a^1_\theta\|^2,
\end{aligned} \tag{C.19}$$

where the equality follows because $g_\theta(\mu^1) = \bar{r}(\theta, a^1_\theta) - (\mu^1)^\top \bar{y}(\theta, a^1_\theta)$, because $\nabla\bar{r}(\theta, a^1_\theta) = \nabla\bar{y}(\theta, a^1_\theta)^\top \mu^1$ from the first order condition of g_θ (by assumption CA 4, a^1_θ is interior), and because $\nabla g_\theta(\mu^1) = -\bar{y}(\theta, a^1_\theta)$ from the envelope theorem applied to g_θ (using compactness of $\mathcal{A}$ and Assumptions CA 1-CA 3 we can apply Corollary 4 in [108]).

We now proceed to bound $g_\theta(\mu)$. Fix $\mu \in \mathcal{N}(\mu^1, \nu)$. We have

$$\begin{aligned}
g_\theta(\mu) &= \max_{a \in \mathcal{A}} \left\{\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\right\} \geq \max_{\{a : \|a - a^1_\theta\| \leq \varphi\}} \left\{\bar{r}(\theta, a) - \mu^\top \bar{y}(\theta, a)\right\} \\
&\geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top (\mu - \mu^1) + \max_{\{x : \|x\| \leq \varphi\}} \left\{\left(\nabla\bar{y}(\theta, a^1_\theta)^\top (\mu^1 - \mu)\right)^\top x - \frac{\kappa}{2}\|x\|^2\right\} \\
&= g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top (\mu - \mu^1) + h\left(\nabla\bar{y}(\theta, a^1_\theta)^\top (\mu^1 - \mu)\right),
\end{aligned}$$

where the first inequality follows from restricting the optimization to $a \in \mathcal{A}$ such that $\|a - a^1_\theta\| \leq \varphi$; the second from (C.19), using Cauchy-Schwartz and the triangle inequality to bound $\mu^\top \kappa_y \leq \|\mu\|\|\kappa_y\| \leq (\|\mu - \mu^1\| + \|\mu^1\|)\|\kappa_y\| \leq (\nu + \|\mu^1\|)\|\kappa_y\|$, setting $\kappa = \kappa_r + (\nu + \|\mu^1\|)\|\kappa_y\|$, and making the change of variables $x = a - a^1_\theta$; the second equality follows from setting $h(z) = \max_{\{x : \|x\| \leq \varphi\}} \left\{z^\top x - \frac{\kappa}{2}\|x\|^2\right\}$.

Note that $h(z)$ is the convex conjugate of $\kappa f(x)$ with $f(x)$ defined in the statement of Lemma 18. Using that the convex conjugate of the scaled function $\kappa f(x)$ is given by

$\kappa f^*(z/\kappa)$ (see, e.g., [30, Section 3.3.2]) together with Lemma 18 and that the dual norm to the Euclidean norm is the Euclidean norm, we obtain that

$$h(z) = \begin{cases} \frac{1}{2\kappa}\|z\|^2 & \text{if } \|z\| \leq \kappa\varphi \\ \varphi\|z\| - \frac{1}{2}\kappa\varphi^2 & \text{otherwise.} \end{cases}$$

Given that σ is a lower bound on the smallest singular value of $\nabla\bar{y}(\theta, a_\theta^1)$, it holds that $\|\nabla\bar{y}(\theta, a_\theta^1)^\top z\| \geq \sigma\|z\|$ for all z (see [76], Lemma 3.3). We can equivalently write $h(z) = \frac{1}{\kappa}\min(\kappa\varphi, \|z\|)\cdot\|z\| - \frac{1}{2\kappa}\min(\kappa\varphi, \|z\|)^2$, which implies that $h(\tilde{z}) \geq h(z)$ whenever $\|\tilde{z}\| \geq \|z\|$ since h is increasing in $\|z\|$. This yields that $h(\nabla\bar{y}(\theta, a_\theta^1)^\top z) \geq h(\sigma z)$ for all $z \in \mathbb{R}^L$.

Therefore, if $\mu \in \mathcal{N}(\mu^1, \nu)$, we have that $\|\mu - \mu^1\| \leq \nu = \kappa\varphi/\sigma$, in which case $h(\sigma z) = (\sigma^2/2\kappa)\|z\|^2$, which yields

$$g_\theta(\mu) \geq g_\theta(\mu^1) + \nabla g_\theta(\mu^1)^\top(\mu - \mu^1) + \frac{\sigma^2}{2\kappa}\|\mu - \mu^1\|^2,$$

and the result follows.

□

C.3.4 Proof of Lemma 9

Let us prove that the first statement of Assumption 2 holds by showing that the function $J(\cdot)$ is linear over the set $\mathcal{N}(\rho_1, \delta) = \{\rho : \|\rho - \rho_1\| \leq \delta\}$ with δ given in the statement of this result.

Take $\rho \in \mathcal{N}(\rho_1, \delta)$. That is, $\rho = \rho_1 + \varepsilon v$ for some v unitary vector and ε a positive real number smaller than δ. Let $\xi^\top = (\rho_1, \mathbf{1})$ be the corresponding right hand side of problem (C.2) for $\rho = \rho_1$ and $u^\top = (v, \mathbf{0})$, where $\mathbf{1}$ and $\mathbf{0}$ denote a vector of ones and zeros, with a proper size, respectively. Note that in this case, $B^{-1}u = B_{\rho_1}^{-1}v$ because the last $|\Theta|$ components of u are zero and thus $\|B^{-1}u\| = \|B_{\rho_1}^{-1}v\| \leq \|B_{\rho_1}^{-1}\|$ where the last equality hold because v is an unitary vector. Then, by Lemma 19 we have that if $0 \leq \varepsilon \leq \frac{\phi^*_{\min}}{\|B_{\rho_1}^{-1}\|}$, B is an optimal basis for the standard problem with right hand side $\xi + \varepsilon u$ and therefore the optimal basic variable vector, namely ϕ_B, can be computed as $B^{-1}(\xi + \varepsilon u)$. Let us define c the objective function coefficient vector of problem (C.2). That is, $c_{\mathrm{i}} = p_{\mathrm{i}}\bar{r}_{\mathrm{i}}$ if $\mathrm{i} \in \Theta$ and 0 otherwise. Thus, calling c_B to the coefficient vector associated to the basic variables, it holds that

$$\begin{aligned} J(\rho) &= c_B^\top B^{-1}(\xi + \varepsilon u) \\ &= c_B^\top B^{-1}\xi + \varepsilon c_B^\top B^{-1}u \\ &= c_B^\top B^{-1}\xi + c_B^\top B_{\rho_1}^{-1}(\varepsilon v) \\ &= J(\rho_1) + \nabla J(\rho_1)(\rho - \rho_1)\,, \end{aligned}$$

where the last equality follows because B is optimal basis of problem (C.2) and $\varepsilon v = \rho - \rho_1$. We then have that $J(\cdot)$ is linear over $\mathcal{N}(\rho_1, \delta)$ and the first statement holds with $K = 0$.

Note that the second statement follows directly because, by hypothesis, the constraints are binding for ρ_1 and taking $\rho \in \mathcal{N}(\rho_1, \delta)$, by Lemma 19- applied to $\xi + \varepsilon u$ with $\varepsilon \leq \delta, u^\top = (v, \mathbf{0})$ and $\|v\| = 1$- holds that the optimal basis does not change.

Therefore, we conclude that Assumption 2 holds for $\delta = \frac{\phi^*_{\min}}{\|B^{-1}_{\rho_1}\|}$, and $K = 0$.

$\square$

C.4 Proofs for Section 3.8

C.4.1 Proof of Proposition 9

Recall that there exist μ^* optimal dual solution satisfying $\mu^* \in [0, \bar{\mu}]$ (see Step 1 in the proof of Proposition 7). Thus, it is enough to show that $\left(\beta\left(\frac{\theta}{1+\mu^*}\right), \mu^*\right)$ is an optimal solution-Lagrange multiplier pair. We will proceed by using Proposition 5.1.5 in [23]. That is, we need to check primal and dual feasibility, Lagrangian optimality and complementary slackness.

1. *Dual feasibility.* It follows directly because we take μ^* optimal dual solution.

2. *Primal feasibility and complementary slackness.* To check primal feasibility and complementary slackness we will apply Proposition 2.1.2 in [23], which gives us that, as μ^* is optimal dual solution, we have that

$$\Psi'_\rho(\mu^*)(\mu - \mu^*) \geq 0, \ \forall \mu \in [0, \bar{\mu}],$$

where the derivative of Ψ_ρ is given by

$$\Psi'_\rho(\mu) = \rho + \sum_{\theta \in \Theta} p_\theta g'_\theta(\mu) = \rho - \sum_{\theta \in \Theta} p_\theta \bar{m}\left(\beta\left(\frac{\theta}{1+\mu}\right)\right). \tag{C.20}$$

If $\mu^* = 0, \Psi'_\rho(0)\mu \geq 0$ and therefore $\Psi'_\rho(0) \geq 0$. Note that we also have $\Psi_\rho(\mu^*)\mu^* = 0$, and then primal feasibility and complementary slackness follows by (C.20) because $\sum_{\theta \in \Theta} p_\theta \bar{m}\left(\beta\left(\frac{\theta}{1+\mu^*}\right)\right)$ is the expected payment under the optimal bidding strategy.

If $\mu^* > 0$, there exists $\nu > 0$ such that $\mu^* + \nu$ and $\mu^* - \nu$ belongs to $[0, \bar{\mu}]$. Therefore both $\Psi'_\rho(\mu^*)\nu$ and $\Psi'_\rho(\mu^*)(-\nu)$ and non-negative, obtaining $\Psi'_\rho(\mu^*) = 0$, and primal feasibility and complementary slackness hold.

3. *Lagrangian optimality.* Note that

$$\begin{aligned}
\arg\max_{\phi\in\Phi} \mathcal{L}(\phi,\mu^*) &= \arg\max_{\phi\in\Phi} \left\{ \mu^*\rho + \sum_{\theta\in\Theta} p_\theta \int_{\mathcal{A}} (\theta\bar{q}(a) - (1+\mu^*)\bar{m}(a))\, \mathrm{d}\phi_\theta(a) \right\} \\
&= \left\{ \arg\max_{\phi_\theta\in\Delta(\mathcal{A})} \int_{\mathcal{A}} (\theta\bar{q}(a) - (1+\mu^*)\bar{m}(a))\, \mathrm{d}\phi_\theta(a) \right\}_{\theta\in\Theta} \\
&= \left\{ \arg\max_{a\in\mathcal{A}} g_\theta(a,\mu^*) \right\}_{\theta\in\Theta},
\end{aligned}$$

where the second equality holds because we can separate the problem for each θ. But note that $g_\theta(a,\mu^*)$ is maximized at $a = \beta(\theta/\mu^* + 1)$ and thus we have Lagrangian optimality.

Therefore, the four conditions holds and the proof is completed. □

C.5 Auxiliary Results

The following lemma is a technical result we need to prove Lemma 16.

Lemma 17. *For every* $a, b \in \mathbb{R}^+$

$$\int_0^T a\exp(-b(T-t)) \wedge 1\, \mathrm{d}t \leq \frac{1}{b}(\log(a) + 1).$$

Proof. Let $\tilde{T}$ be the real number such that $a\exp(-b(T-\tilde{T})) = 1$. Then, by doing some math we obtain

$$\begin{aligned}
\int_0^T a\exp(-b(T-t)) \wedge 1\, \mathrm{d}t &= \int_0^{\tilde{T}} a\exp(-b(T-t))\, \mathrm{d}t + T - \tilde{T} \\
&= \frac{a}{b}\exp(-b(T-t))\Big|_0^{\tilde{T}} + T - \tilde{T} \\
&= \frac{1}{b} a\exp(-b(T-\tilde{T})) - \frac{a}{b}\exp(-bT) + T - \tilde{T} \\
&\leq \frac{1}{b} + T - \tilde{T}.
\end{aligned}$$

On the other hand, as $\exp(-b(T-\tilde{T})) = 1$ we have that $T - \tilde{T} = \frac{\log a}{b}$, and therefore we conclude that

$$\int_0^T a\exp(-b(T-t)) \wedge 1\, \mathrm{d}t \leq \frac{1}{b}(\log a + 1).$$ □

Lemma 18. *Let $\|x\|$ be a norm in the Euclidean space and let $\|z\|_* = \max_{\|x\|\leq 1}\{z^\top x\}$ be its dual norm. Let $f(x) = \frac{1}{2}\|x\|^2$ if $\|x\| \leq \varphi$ and $f(x) = \infty$ otherwise. Then, its convex conjugate $f^*(z) = \max_x \left\{z^\top x - f(x)\right\} = \max_{x:\|x\|\leq\varphi} \left\{z^\top x - \frac{1}{2}\|x\|^2\right\}$ is given by*

$$f^*(z) = \begin{cases} \frac{1}{2}\|z\|_*^2 & \textit{if } \|z\|_* \leq \varphi \\ \varphi\|z\|_* - \frac{1}{2}\varphi^2 & \textit{otherwise.} \end{cases}$$

Proof. Note that the convex conjugate can be more compactly written as $\min(\varphi, \|z\|_*) \cdot \|z\|_* - \frac{1}{2}\min(\varphi, \|z\|_*)^2$. We first show that the latter expression provides an upper bound and then show that the upper can be attained by choosing a suitable feasible solution.

For the upper bound, use Cauchy-Schwartz inequality to obtain that

$$f^*(z) \leq \max_{x:\|x\|\leq\varphi} \left\{\|z\|_*\|x\| - \frac{1}{2}\|x\|^2\right\} = \max_{\ell\in\mathbb{R}:0\leq\ell\leq\varphi} \left\{\|z\|_*\ell - \frac{1}{2}\ell^2\right\},$$

where the equality follows because we can equivalently optimize over the attainable norm values in $[0, \varphi]$. The objective value of the latter problem is a downward parabola with maximum at $\ell = \|z\|_*$. The claim follows because the optimal solution is $\ell = \min(\varphi, \|z\|_*)$.

For the lower bound, fix z and let $\tilde{x} = \arg\max_{\|x\|\leq 1}\{z^\top x\}$, i.e., a vector satisfying $\|z\|_* = z^\top\tilde{x}$. Such a vector exists because the dual norm always admits an optimal solution by Weierstrass theorem (the objective is continuous and the feasible set is compact). Consider the solution $x = \min(\varphi, \|z\|_*)\tilde{x}$. This solution is feasible because $\|x\| = \min(\varphi, \|z\|_*)\|\tilde{x}\| \leq \varphi$ since $\|\tilde{x}\| \leq 1$. Therefore,

$$\begin{aligned} f^*(z) &\geq z^\top x - \frac{1}{2}\|x\|^2 = z^\top\tilde{x} \cdot \min(\varphi, \|z\|_*) - \frac{1}{2}\|\tilde{x}\|^2 \cdot \min(\varphi, \|z\|_*)^2 \\ &\geq \min(\varphi, \|z\|_*) \cdot \|z\|_* - \frac{1}{2}\min(\varphi, \|z\|_*)^2, \end{aligned}$$

where the last inequality follows because $\|z\|_* = z^\top\tilde{x}$ and $\tilde{x} \leq 1$. The result follows. □

Lemma 19. *Consider the general linear program problem*

$$\begin{aligned} \max_x \ & c^\top x \\ \text{s.t } & Ax = \xi + \varepsilon u, \\ & x \geq 0, \end{aligned} \tag{C.21}$$

where $c, x, \mathbf{0}, \mathbf{1} \in \mathbb{R}^n$, $\xi, u \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ matrix of rank m and ε a real parameter. Define $x^*_{\text{min}} = \min\{x^*_{\text{i}} : x^*_{\text{i}} > 0\}$, where x^* denotes a non-degenerate optimal solution of problem (C.21) for $\varepsilon = 0$, and denote by $A_B \in \mathbb{R}^{m\times m}$ its associated basis matrix. If $\delta = x^*_{\text{min}}/\|A_B^{-1}u\|$, then A_B remains optimal for problem (C.21) for all $0 \leq \varepsilon \leq \delta$.

Proof. By permuting its columns, matrix A can be written as $A = (A_B|A_N)$, where $A_B \in \mathbb{R}^{m\times m}$ is the submatrix containing the columns associated to the basic variables of x^* and $A_N \in \mathbb{R}^{m\times(n-m)}$ is the submatrix corresponding to the non-basic variables of x^*. Furthermore, we can write $x^* = (x^*_B, \mathbf{0})$, where $x^*_B = A_B^{-1}\xi \in \mathbb{R}^m$ is the subvector of basic variables and $\mathbf{0} \in \mathbb{R}^{n-m}$. Note that non-degeneracy of x^* implies that $x^*_B > 0$.

Note that $\delta > 0$ is well defined because $\{x^*_{\mathrm{i}} : x^*_{\mathrm{i}} > 0\}$ is not empty due to the non-degeneracy condition on x^*. Take $\varepsilon \leq \delta$. We will prove that A_B is an optimal basis for (C.21), that is, $x = (x_B, \mathbf{0})$ with $x_B = A_B^{-1}(\xi+\varepsilon u)$ is an optimal solution for Problem (C.21). Changing the right-hand side of the equality constraints does not change the reduced cost vector and, therefore, it is enough to show that x_B is non-negative.

To this end, take $j \in \{1, \ldots, m\}$ such that $(A_B^{-1}u)_j < 0$. Note that if does not exist such j, the desired inequality follows trivially because $x^*_j = (A_B^{-1}\xi)_j > 0$ since x^* is non-degenerate. Otherwise, we have that

$$(x_B)_j = \left(A_B^{-1}(\xi + \varepsilon u)\right)_j = (x^*_B)_j + \varepsilon \left(A_B^{-1}u\right)_j \geq x^*_{\min} - \varepsilon\|A_B^{-1}u\| \geq x^*_{\min} - \delta\|A_B^{-1}u\| = 0\,,$$

where the first equation follows from the definition of x_B, the second because $(x^*_B)_j$ is a basic variable, the first inequality from the definition of $x^*_{\min}$ together with $|x_j| \leq \|x\|$ for every $x_j \in \mathbb{R}^m$, the second inequality because $\varepsilon \leq \delta$, and the last from the definition of δ. The proof is completed. □