The FedEx Problem

Amos Fiat [*] Kira Goldner[†] Anna R. Karlin[‡] Elias Koutsoupias[§]

November 5, 2020

Abstract

"Remember that Time is Money"

— Benjamin Franklin in *Advice to a Young Tradesman* (1748)

Consider the following setting: a customer has a package and is willing to pay up to some value v to ship it, but needs it to be shipped by some deadline d. Given the joint prior distribution from which (v, d) pairs are drawn, we characterize the auction that yields optimal revenue, contributing to the limited understanding of optimal auctions beyond single-parameter settings. Our work requires a new way of combining and ironing revenue curves which illustrate why randomization is necessary to achieve optimal revenue. Finally, we strengthen the emerging understanding that duality is useful for both the design and analysis of optimal auctions in multi-parameter settings.

1 Introduction

Consider the pricing problem faced by FedEx. Each of their customers has a deadline d by which he needs his package to arrive, and a value v for receiving the package by the deadline. The customer's utility for getting his package shipped by day i at a price of p is $v - p$ if $i \leq d$ (i.e., it is received by his deadline) and $-p$ otherwise. Of course, a customer's (v, d) pair is the private information of the customer. We study the Bayesian setting, where this pair (v, d) is drawn from a prior distribution known to FedEx, and address the question of optimal (revenue maximizing) mechanism design. Note that the prior distribution may be arbitrarily correlated.

Suppose that FedEx offers a discrete set of shipping options (1-day, 2-day, 3-day, up to m-day shipping). The prior that FedEx has on its customer's needs is given by a probability distribution $(q_1, q_2, \dots, q_m)$, where q_i is the probability that the customer has a deadline i days from now, and a set of marginal value distributions, where F_i, for $1 \leq i \leq m$, is the distribution of values given that the customer's deadline is i.

[*]Tel Aviv University. fiat@tau.ac.il.

[†]University of Washington. kgoldner@cs.washington.edu. Funded by the National Science Foundation under CCF grant 1420381.

[‡]University of Washington. karlin@cs.washington.edu. Funded by the National Science Foundation under CCF grant 1420381.

[§]University of Oxford. elias@cs.ox.ac.uk. Funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 321171 (ALGAME).

This research was done in part while the authors were visiting the Simons Institute for Theoretical Computer Science.

We consider the single agent problem in this paper, or equivalently, the setting where FedEx sells to identically drawn buyers and has constraints on the number of buyers it can supply. We obtain a closed form, clean, and optimal auction for this setting. Our work adds to the relatively short list of multi-parameter settings where a closed form solution is known. (See related work below for more on this.)

The pricing problem we consider is extremely natural and arises in numerous scenarios, whether it is Amazon.com providing shipping options, Internet Service Providers offering bandwidth plans, or a myriad of other settings in which a seller can price discriminate or otherwise segment her market by delaying service, or providing lower quality/cheaper versions of a product. In particular, this setting is relevant whenever a customer has a value and a sensitivity to time or some other feature of service. A "deadline" represents the base level of need, imposed on a buyer by outside circumstances, whereas a valuation represents the buyer's own willingness to pay. It is important to understand how buyer deadline constraints impact the design of auctions and what leverage they give to the auctioneer to extract more revenue.

1.1 Related Work

The FedEx setting is a variant of price discrimination in which the customers are grouped by their deadline. Price discrimination offers different prices to users with the goal of improving revenue [Bergemann et al., 2015]. Alternatively one can view the FedEx problem as a multi-dimensional optimal auction problem. There are two ways to express the FedEx problem in this way. First, as a 2-dimensional (value $\times$ deadline) problem of arbitrary joint distribution in which the second variable takes only integer values in a bounded interval. Alternatively, as a very special case of the m-dimensional unit-demand problem with correlated values (the customer buys a shipping option among the m choices)—his value for the first d options is v, and for the last $m - d$ is 0.

There is an extensive body of literature on optimal auction design. The seminal work of Myerson [1981] has completely settled the case of selling a single item to multiple bidders and extends directly to the more general framework of single-parameter settings. Note that Riley and Zeckhauser [1983] also prove that the optimal single-parameter mechanism is deterministic, and that Bulow and Roberts [1989] are responsible for the interpretation of virtual values as a marginal contribution to revenue.

The most complicated part of Myerson's solution is his handling of distributions that are not regular by "ironing" them, that is, by replacing the revenue curves by their upper concave envelope. Myerson's ironing is done in quantile space. In this work, we also need to iron the revenue curves, but we need to do this in value space.

Extending Myerson's solution to the multi-dimensional case has been one of the most important open problems in Microeconomics. For the case of unit-demand agents, a beautiful sequence of papers [Chawla et al., 2007; Briest et al., 2015; Chawla et al., 2015, 2010; Alaei, 2011; Cai and Daskalakis, 2011] showed how to obtain approximately optimal auctions. For the case of finite type spaces, [Cai et al., 2012, 2013b,a] are able to use linear and convex programming techniques to formulate and solve the optimal auction problem. This gives a black-box reduction from mechanism to algorithm design that yields a PTAS for revenue maximization in unit-demand settings. For the case of additive agents, additional recent breakthroughs [Hart and Nisan, 2017; Li and Yao, 2013; Babaioff et al., 2014; Yao, 2015; Cai et al., 2016; Cai and Huang, 2013] have also resulted in approximately optimal mechanisms.

But if we insist on optimal auctions for continuous probability distributions, no general solution is known for the multi-dimensional case—even for the two-dimensional single-bidder case—and it is very possible that no such simple solution exists for the general case. One of the reasons that the multi-dimensional case is so complex is that optimal auctions are not necessarily deterministic [Pavlov, 2011; Thanassoulis, 2004; Briest et al., 2015; Hart and Reny, 2012; Hart and Nisan, 2013; Manelli and Vincent, 2006; Pycia, 2006; Daskalakis et al., 2013]. The optimal auction for the FedEx problem also turns out to be randomized with exponentially many different price levels in the worst case [Saxena et al., 2018].

There are some relevant results that solve special cases of the two-parameter setting. One of the earliest works is by Laffont et al. [1987] who study a distant variant of the FedEx problem. In their problem, the bidder has two parameters, a and b, each uniformly distributed on $[0, 1]$, and the bidder's utility function is very specifically the quadratic function $ax - \frac{1}{2}(b+1)x^2 - p$. Here, x is a single-dimensional allocation variable and p is the payment variable. The idea is that both the slope and the intercept of the buyer's demand curve are unknown to the seller. To solve this problem, Laffont et al. come up with a change of variable technique to use only one variable when solving for the allocation in both parameters. By solving the optimization problem and the resulting integration by hand with this technique, they provide a highly non-trivial closed-form allocation rule, demonstrating that even the simplest independent two-parameter settings are far more difficult than single-parameter settings.

McAfee and McMillan [1988] study a generalization of this problem. First, they characterize incentive-compatibility precisely in direct, deterministic, and differentiable mechanisms. Then, they reference a notion of "single-crossing" which says that the marginal rate of substitution[1], must be monotone increasing in the buyer's type. McAfee and McMillan generalize this condition to multi-parameter settings, and then extend the analysis of Laffont et al. (using the same change of variables technique) to any number of variables if they satisfy generalized single-crossing and other small conditions. However, their analysis only applies when the optimal mechanism is deterministic. Finally, McAfee and McMillan also consider the setting where a buyer has independent valuations for m heterogenous items. They prove that under a "regularity" condition[2], for $m = 2$, the optimal mechanism is deterministic, and they further reason from prior results that the optimal mechanism would set a price for each item individually as well as the grand bundle.

These initial results were followed by more general results. In particular, Haghpanah and Hartline [2015] consider the problem of selling a product with multiple quality levels to unit-demand bidders. The mechanism they consider is selling only the highest quality product at a posted price. (In the FedEx problem, this corresponds to having a single price for every shipping option.) When the buyers' value distributions have a specific type of positive correlation, then this mechanism is optimal, because the high-valued customers are less quality-sensitive, and thus will not pay a premium for a different outcome. Haghpanah and Hartline solve for when this mechanism is a point-wise virtual value maximizer, with expected revenue equal to virtual welfare, and then solve for the paths of tight IC constraints to integrate over, effectively reverse engineering the virtual value functions. This approach also corresponds to proving when bundling is optimal in particular additive settings. Their work generalizes results from Armstrong [1996].

Daskalakis et al. [2017] establish a duality framework where the primal is expressed in terms of utility and a transformed measure μ of the buyer distributions, and the dual is an optimal

[1] This is equal to the derivative of utility with respect to the allocation divided by the negative derivative of utility with respect to payment. This also the negative shadow price.

[2] They assume that $t \cdot f'(t) + (m+1)f(t) \geq 0$. Note that this is *not* Myerson's regularity condition, but would be if the $(m+1)$ were replaced by 1.

transport problem. The dual variables are a measure μ' that stochastically dominate the primal measure μ, and the objective is the distance between the positive part of the measure μ'_+ and the negative part of the measure μ'_-. First, Daskalakis et al. establish that strong duality holds in their duality framework, so the dual can be used to solve for or certify the optimal primal. Then, using their duality framework, for any mechanism with a finite menu size (number of outcomes), they give a characterization in terms of the stochastic dominance conditions on the measure μ. One application of this result proves that for m items distributed i.i.d. on $[c, c+1]$ for large enough c, grand bundling is optimal.

Our approach is based on a duality framework. Two such frameworks have been proposed. The first framework by Daskalakis et al. [2017] expresses the primal in terms of utility and a transformed measure μ of the buyer distributions, and the dual is an optimal transport problem. The dual variables are a measure μ' that stochastically dominate the primal measure μ, and the objective is the distance between the positive part of the measure μ'_+ and the negative part of the measure μ'_-. First, Daskalakis et al. establish that strong duality holds in their duality framework, so the dual can be used to solve for or certify the optimal primal. Then, using their duality framework, for any mechanism with a finite menu size (number of outcomes), they give a characterization in terms of the stochastic dominance conditions on the measure μ. One application of this result proves that for $m-1$ items distributed i.i.d. on $[c, c+1]$ for large enough c, grand bundling is optimal.

The second framework is by Giannakopoulos and Koutsoupias [2014], which is based on expressing the revenue maximization problem as an optimization problem in terms of utility functions and their partial derivatives. They find primal and dual variables that are both feasible and also satisfy complementary slackness, and thus via weak duality and complementarity, have equal objectives. Using their framework, they prove that the Straight-Jacket-Auction[3] is optimal for an additive bidder whose item valuations are i.i.d. from $U[0, 1]$ for up to six items. In [Giannakopoulos and Koutsoupias, 2015], the authors subsequently use their framework to give closed-form optimal allocation and payment rules for several independent non-identical two-item problems (where the distributions are from monomial or exponential families over $[0, H]$), and the mechanisms are no longer deterministic. The duality framework of Giannakopoulos and Koutsoupias is a fairly general approach, but their applications still require fairly strong assumptions about the distributions in order to make progress on characterizing optimal auctions. Our solution of the FedEx problem follows this latter duality framework.

For much more on both exact and approximate optimal mechanism design, see [Daskalakis, 2015; Chawla and Sivan, 2014; Roughgarden, 2015; Hartline, 2013; Cai et al., 2011]. For background on duality in infinite linear and convex programs, see e.g., [Anderson and Nash, 1987; Luenberger, 1997].

1.2 Our Contribution

Our result is one of the first explicit closed-form generalizations of [Myerson, 1981] to multi-parameter settings with arbitrary (joint) distributions, and contributes to recent breakthroughs in this space. We use a duality framework where we prove optimality by finding primal and

[3] The Straight-Jacket-Auction for m items is a deterministic mechanism where a price is set for every bundle size, bound by sale probabilities. The price for the bundle of size r, $p_r^{(m)}$, is such that, given prices $p_1^{(m)}$ through $p_{r-1}^{(m)}$ already set, a buyer with value $v \sim [0, 1]^r \times 0^{m-r}$ will not buy any bundle of size r or smaller with probability $1 - \frac{r}{m+1}$.

dual solutions that satisfy sufficient conditions. The optimal primal and dual variables have an interesting inductive structure, and the allocation rule is potentially randomized over at most 2^{i-1} prices on i-day. Our approach strengthens the emerging understanding that duality is useful for determining the structure of the optimal auction in non-trivial settings, in addition to its use in analyzing the auction.

In Myerson's setting, the "ironing" of revenue curves and virtual valuations to determine the optimal auction is required to enforce incentive compatibility constraints among multiple bidders. In our setting, we need a form of ironing even for one bidder in order to enforce incentive compatibility constraints among the multiple options. This work also suggests that ironing is one of the biggest hurdles in extending Myerson to more general settings.

2 Preliminaries

As discussed above, the type of a customer is a (value, deadline) pair. An auction takes as input a reported type $t = (v, d)$ and determines the shipping date in $\{1, \ldots, m\}$ and the price. We denote by $a_i(v)$ the probability that the package is shipped by day i when the agent reports (v, i), and by $p_i(v)$ the corresponding expected payment (the expectation is taken over the randomness in the mechanism).

Our goal is to design an optimal mechanism for this setting. By the revelation principle, we can restrict our attention to incentive compatible mechanisms. In this setting, when an agent with type (v, i) reports a type of (v', i'), he has utility

$$u(v', i' \mid v, i) = \begin{cases} va_{i'}(v') - p_{i'}(v') & \text{if } i' \leq i \\ -p_{i'}(v') & \text{otherwise.} \end{cases}$$

The incentive compatibility requirement is that

$$u(v, i) \geq u(v', i' \mid v, i) \qquad \forall v', i'. \tag{1}$$

We also require individual rationality, i.e., $u(v, i) \geq 0$ for all (v, i). Without loss of generality, $a_i(v)$ is the probability that the package is delivered *on day* i, since any incentive compatible mechanism which delivers a package early can be converted to one that always delivers on the deadline, while retaining incentive compatibility and without losing any revenue.

For each fixed i, this implies the standard (single parameter) constraints [Myerson, 1981], namely

$$\forall i,\ a_i(v) \text{ is monotone weakly increasing and in } [0, 1]; \tag{2}$$

$$\forall i,\ p_i(v) = va_i(v) - \int_0^v a_i(x)dx \quad \text{and hence} \quad u(v, i) = \int_0^v a_i(x)dx. \tag{3}$$

Clearly no agent would ever report $i' > i$, as this would result in non-positive utility. However, we do need to make sure that the agent has no incentive to report an earlier deadline, and hence another IC constraint is that for all $2 \leq i \leq m$:

$$u(v, i-1|v, i) \leq u(v, i) \tag{4}$$

which is equivalent to

$$\int_0^v a_{i-1}(x)dx \leq \int_0^v a_i(x)dx \qquad \forall i \text{ s.t. } 1 < i \leq m. \tag{5}$$

We sometimes refer to this as the inter-day IC constraint. Since $a_i(v)$ is the probability of allocation of i-day shipping given report (v, i), constraints (2), (3) and (5) are necessary and sufficient, by transitivity, to ensure that

$$u(v, i) \geq u(v', i' | v, i)$$

for *all* possible misreports (v', i').

The prior. We assume that the agent's (value, deadline) comes from a known joint prior distribution F. Let q_i be the probability that the customer has a deadline $i \in \{1, \ldots, m\}$, that is,

$$q_i = \Pr_{(v,d)\sim F}[d = i]$$

and let $F_i(\cdot)$ be the marginal distribution function of values for bidders with deadline i. That is,

$$F_i(x) = \Pr_{(v,d)\sim F}[v \leq x \mid d = i].$$

We assume that F_i is atomless and strictly increasing, with density function defined on $[0, H]$. Let $f_i(v)$ be the derivative of $F_i(v)$.

The objective. Let $\varphi_i(v) = v - \frac{1-F_i(v)}{f_i(v)}$ be the virtual value function for v drawn from distribution F_i. Applying the Myerson payment identity (3) implies that the expected payment of a customer with deadline i is

$$\mathbb{E}_{v\sim F_i}[p_i(v)] = \mathbb{E}_{v\sim F_i}[\varphi_i(v)a_i(v)].$$

Thus, we wish to choose monotone allocation rules $a_i(v)$ for days $1 \leq i \leq m$, so as to maximize

$$\mathbb{E}_{(v,i)\sim F}[p_i(v)] = \sum_{i=1}^m q_i \mathbb{E}_{v\sim F_i}[p_i(v)] = \sum_{i=1}^m q_i \mathbb{E}_{v\sim F_i}[\varphi_i(v)a_i(v)] = \sum_{i=1}^m q_i \int_0^H \varphi_i(v) f_i(v) a_i(v) dv,$$

subject to (2) and (5).

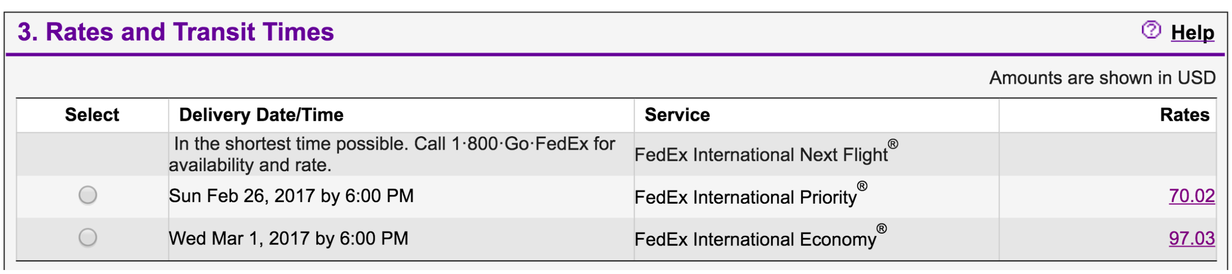
A trivial case and discussion. If we knew that the customer that would arrive would have deadline i and we could thus condition on this event, ensuring that his value is drawn from the marginal distribution F_i, then the optimal pricing would be trivial, as this is a single-agent, single-item auction. In this case, the optimal mechanism for such a customer is to set the price for service by day i to the reserve price r_i for his prior. If we just had a number of single-dimensional problems, one for each deadline, we would want to set a price of r_i for each i-day shipping option. If it is the case that $r_i \geq r_{i+1}$ for each i, then the entire Fedex problem is trivial, since setting r_i as the price for i-day shipping satisfies all of the IC constraints, and this pointwise optimizes each conditional distribution.

Note that even should the marginal distribution for buyer values for 1-day shipping stochastically dominate the marginal distribution for 2-day shipping and so on, the later shipping options may still have higher reserve prices. For example, if F_2 is uniform over the set $\{1, 10\}$, the reserve is 10. If F_1 is uniform over the set $\{9, 10\}$, the reserve is 9. Hence, we do not make the assumption that the reserve prices are weakly decreasing with the deadline.

In fact, we do not make the assumption of stochastic domination either in order to be fully general. The prior F captures the result of random draws from a population consisting of a mixture of different types. Obviously any particular individual with deadline i is at least as happy with day $i-1$ service as with day i service, but two random individuals may have completely uncorrelated needs. To give an example, an individual ordering a last minute birthday present may have a lower value than an individual scheduling the delivery of surgical equipment that is needed to perform open heart surgery in three weeks time. In fact, for more valuable packages, one could imagine that people take the time to plan ahead.

Another factor has to do with costs. It is likely that the cost that FedEx incurs for sending a package within i days is higher than the cost FedEx incurs for sending a package within $i' > i$ days, since in the latter case, for example, FedEx has more flexibility about which of many planes/trucks to put the package on, and even may be able to reduce the total number of plane/truck trips to a particular destination given this flexibility. More generally, in other applications of this problem, the cost of providing lower quality service is lower than the cost of providing higher quality service. Thus, even if reserve prices tend to decrease with i, all bets are off once we consider a customer's value for deadline i conditioned on that value being above the expected cost to FedEx of shipping a package by deadline i for each i.

In this paper, we are not explicitly modeling the costs that FedEx incurs, the optimization problems that it faces, the online nature of the problem, or any limits on FedEx's ability to ship packages. These are interesting problems for future research. The discussion in the preceding few paragraphs is here merely to explain why the problem remains interesting and relevant even in the with distributions that do not have decreasing reserve prices in the deadlines. Further, note that in Figure 1, FedEx actually did post a larger price for a later shipping option, implying that they estimate the underlying distributions to have increasing reserve prices.[4]



3. Rates and Transit Times — Help

Amounts are shown in USD

Select	Delivery Date/Time	Service	Rates
	In the shortest time possible. Call 1·800·Go·FedEx for availability and rate.	FedEx International Next Flight®	
○	Sun Feb 26, 2017 by 6:00 PM	FedEx International Priority®	70.02
○	Wed Mar 1, 2017 by 6:00 PM	FedEx International Economy®	97.03

Figure 1: FedEx posted a higher price posted for a later shipping option, implying the underlying distributions do not have decreasing reserve prices in the real world. Note also that these prices are not incentive compatible.

[4]Of course, these prices are not incentive compatible, and the author that purchased shipping when presented with these prices did in fact misreport her deadline.

3 Warm-up: The case of $m = 2$

Suppose that the customer might have a deadline of either one day or two days from now. By the taxation principle, the optimal mechanism is a menu, and in this setting consists of a (potentially randomized) price p_i for having the package delivered i days from now.

Let $R_i(v)$ be the i-day *revenue curve*, that is, $R_i(v) := v \cdot [1 - F_i(v)]$. Let $r_i := \text{argmax}_v R_i(v)$ be the price at which expected revenue from a bidder with value drawn from F_i is maximized, and let $R_i^* := R_i(r_i)$ denote this maximum expected revenue. Since R_i^* is the optimal expected revenue from the agent [Myerson, 1981], conditioned on having a deadline of i, then $q_1 R_1^* + q_2 R_2^*$ is an upper bound on the optimal expected revenue for the two-day FedEx problem. If $r_1 \geq r_2$, then this optimum is indeed achievable by an IC mechanism: just set the 1-day shipping price p_1 at r_1 and the 2-day shipping price p_2 at r_2.

But what if $r_2 > r_1$? In this case, the inter-day IC constraint (5) is violated by this pricing (a customer with $i = 2$ will pretend his deadline is $i = 1$).

Attempt #1: One alternative is to consider the optimal single price mechanism (i.e., $p_1 = p_2 = p$). In this case, the optimal choice is clear:

$$p := \text{argmax}_v \; [q_1 R_1(v) + q_2 R_2(v)], \tag{6}$$

i.e., set the price that maximizes the combined revenue from both days. There are cases where this is optimal, e.g., if both F_1 and F_2 are regular. A proof is given in Subsection 4.2.

Attempt #2: Another auction that retains incentive compatibility, and, in some cases, improves performance is to set the 1-day price p_1 at p and the 2-day price at

$$p_2 := \text{argmax}_{v \leq p} \; R_2(v). \tag{7}$$

However, even if we fix $p_1 = p$, further optimization may be possible if F_2 is not regular.

Attempt #3: Consider the concave hull of $R_2(\cdot)$, i.e., the ironed revenue curve. If $R_2(v)$ is maximized at $r_2 > p$ and $R_2(\cdot)$ is ironed at p, then offering a lottery on 2-day with an expected price of p yields higher expected revenue than offering any deterministic 2-day price of p_2. As we shall see, for this case, this solution is actually optimal. (See Figure 2.)

However, if $p > r_2$, (which is possible if F_1 and F_2 are not regular, even if $r_1 < r_2$), then we will see that the optimal 1-day price is indeed higher than r_2, but not necessarily equal to p.

Attempt #4: If $p > r_2$, set the 1-day price at

$$p_1 := \text{argmax}_{v \geq r_2} \; R_1(v).$$

This should make sense: if we're going to set a 1-day price above r_2, we may as well set the 2-day price at r_2, but in that case, the 2-day curve should not influence the pricing for 1-day (except to set a lower bound for it).

Admittedly, this sounds like a tedious case analysis, and extending this reasoning to three or more days gets much worse. Happily, though, there is a nice, and relatively simple way to put all the above elements together to describe the solution, and then, as we shall see in Section 5, prove its optimality via a clean duality proof.

A solution for $m = 2$. Define $\hat{R}(\cdot)$ to be the concave (ironed) revenue curve corresponding to revenue curve $R(\cdot)$. We define the following combined revenue curve, depicted in Figure 3. Let

$$R_{12}(v) := \begin{cases} q_1 R_1(v) + q_2 \hat{R}_2(v) & v \leq r_2 \\ q_1 R_1(v) + q_2 R_2(r_2) & v > r_2. \end{cases} \tag{8}$$

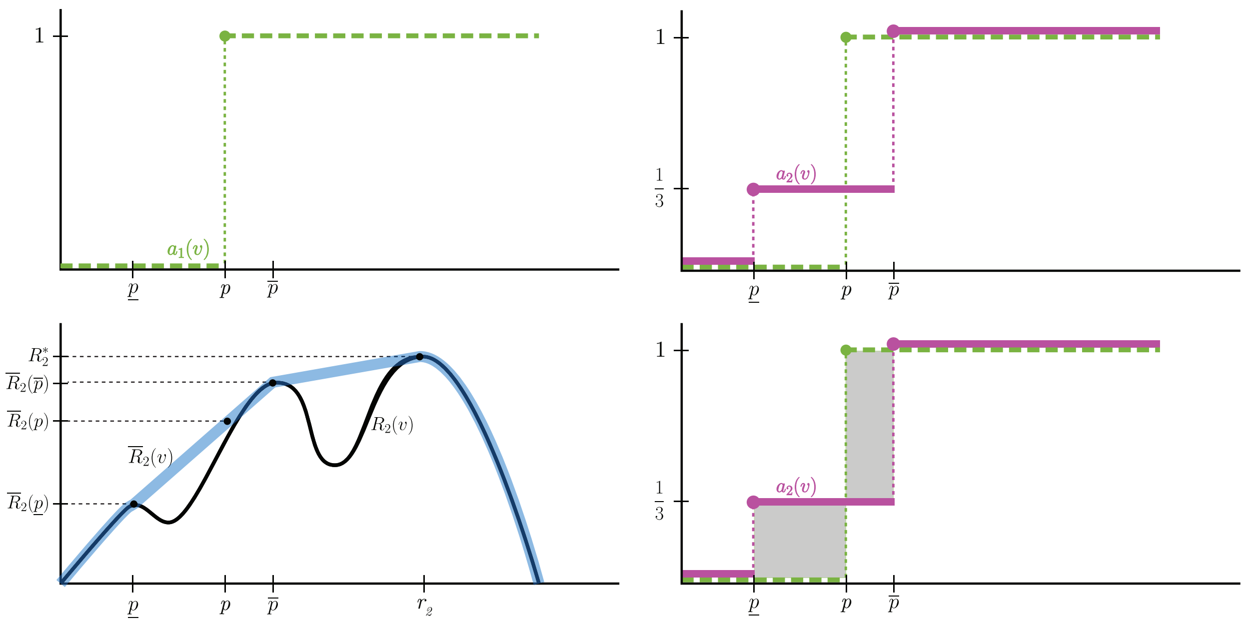


Figure 2: A two-day case: Suppose that the optimal thing to do on 1-day is to offer a price of p. In the upper left, we see the corresponding allocation curve $a_1(v)$. The bottom left graph shows the revenue curve $R_2(\cdot)$ for 2-day (the thin black curve) and the ironed version $\hat{R}_2(\cdot)$ (the thick blue concave curve). Optimizing for 2-day subject to the inter-day IC constraint $\int_0^v a_1(x)dx \leq \int_0^v a_2(x)dx$ suggests that the most revenue we can get from a deadline $d = 2$ customer is $\hat{R}_2(p)$ on 2-day, which can be done by offering the price of $\underline{p}$ with probability 1/3 and a price of $\overline{p}$ with probability 2/3 (since, in this example, $p = (1/3)\underline{p} + (2/3)\overline{p}$). This yields the pink allocation curve $a_2(v)$ shown in the upper right. The fact that these curves satisfy the inter-day IC constraint follows from the fact that the area of the two grey rectangles shown in the bottom right are equal.

Note that because $\hat{R}_2(\cdot)$ is the least concave upper bound on $R_2(\cdot)$ and by definition of r_2 that $\hat{R}_2(r_2) = R_2(r_2)$. The optimal solution is to set

$$p_1 := \text{argmax}_v R_{12}(v),$$

and then take

$$p_2 := r_2 \text{ if } r_2 \leq p_1 \text{ and } \mathbb{E}(p_2) := p_1 \text{ otherwise},$$

where the randomized case is implemented via the lottery as in the example of Figure 2.

The key idea: $R_{12}(v)$ describes the best revenue we can get if we set a price of v for 1-day shipping as shown in Figure 3. Since r_2 is the optimal 2-day price, if we are going to set a price above r_2 for 1-day shipping, then the remaining 2-day optimization problem is unconstrained. On the other hand, if the 1-day price is below r_2, then it would constrain the 2-day price via the inter-day IC constraint (5), and ironing the 2-day revenue curve may be necessary. This is precisely what the definition of $R_{12}(\cdot)$ in (8) does for us. The asymmetry between 1-day and 2-day, specifically the fact that the 1-day curve is never ironed, whereas the 2-day curve is, is a consequence of the inter-day IC constraint (5). We generalize this idea in the next section to solve the m-day problem.

4 An optimal allocation rule

4.1 Preliminaries

As we discussed regarding the objective, our goal is to choose monotone allocation rules $a_i(v)$ for days $1 \leq i \leq m$ so as to maximize $\sum_{i=1}^m q_i \int_0^H \varphi_i(v) f_i(v) a_i(v) dv$.

For a distribution $f_i(\cdot)$ on $[0, H]$ with virtual value function $\varphi_i(\cdot) = v - \frac{1-F_i(v)}{f_i(v)}$, define $\gamma_i(v) := q_i \varphi_i(v) f_i(v)$. Then we aim to choose $a_i(v)$ to maximize $\sum_{i=1}^m \int_0^H \gamma_i(v) a_i(v) dv$.

Let $\Gamma_i(v) = \int_0^v \gamma_i(x) dx$. Observe that this function is the negative of the revenue curve, that is, $\Gamma_i(v) = -q_i R_i(v) = -q_i v[1 - F_i(v)]$. [5] Thus, $\Gamma_i(0) = \Gamma_i(H) = 0$ and $\Gamma_i(v) \leq 0$ for $v \in [0, H]$.

Definition 1. For any function Γ, define $\hat{\Gamma}(\cdot)$ to be the lower convex envelope [6] of $\Gamma(\cdot)$. We say that $\hat{\Gamma}(\cdot)$ is *ironed* at v if $\hat{\Gamma}(v) \neq \Gamma(v)$.

Since $\hat{\Gamma}(\cdot)$ is convex, it is continuously differentiable except at countably many points and its derivative is monotone (weakly) increasing.

Definition 2. Let $\hat{\gamma}(\cdot)$ be the derivative of $\hat{\Gamma}(\cdot)$ and let $\gamma(\cdot)$ be the derivative of $\Gamma(\cdot)$.

Claim 1. The following facts are immediate from the definition of lower convex envelope (See Figure 4.):

- $\hat{\Gamma}(v) \leq \Gamma(v) \quad \forall v.$

[5] $\Gamma_i(v) = q_i \int_0^v [x f_i(x) - (1 - F_i(x))] \, dx$. Integrating the first term by parts gives $\int_0^v x f_i(x) \, dx = v F_i(v) - \int_0^v F_i(x) \, dx$. Combining this with the second term yields $\Gamma_i(v) = -q_i v(1 - F_i(v))$.

[6] The *lower convex envelope* of function $f(x)$ is the supremum over convex functions $g(\cdot)$ such that $g(x) \leq f(x)$ for all x. Notice that the lower convex envelope of $\Gamma(\cdot)$ is the negative of the ironed revenue curve $\hat{R}(v)$.

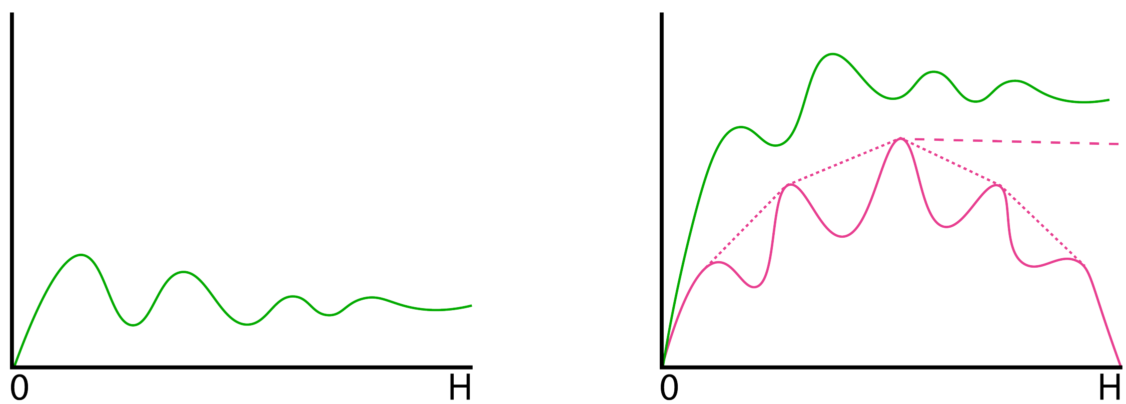


Figure 3: Left: The scaled revenue curve for deadline 1, $q_1R_1(\cdot)$. Right: The pink curve is the scaled revenue curve for deadline 2, $q_2R_2(\cdot)$. The dotted curve represents the ironed $q_2\hat{R}_2(\cdot)$. At any possible 1-day price v, the highest pink point at v is the best revenue that can be obtained from 2-day shipping given that a price of v is set for 1-day. This is either the revenue from the ironed curve $q_2\hat{R}_2(v)$ or, when possible, the revenue of setting a price of r_2, which yields $q_2R_2(r_2)$. The higher green curve is $R_{12}(\cdot)$, the sum of the green curve from the left and the upper pink envelope, which gives the combined revenue from setting a price of v for 1-day and then doing the best thing for 2-day shipping.

- $\hat{\Gamma}(v_{\min}) = \Gamma(v_{\min})$ where $v_{\min} = \text{argmin}_v \Gamma(v)$. (This implies that there is no ironed interval containing $v_{\min}$.)
- $\hat{\gamma}(v)$ is an increasing function of v and hence its derivative $\hat{\gamma}'(v) \geq 0$ is non-negative for all v.
- If $\hat{\Gamma}(v)$ is ironed in the interval $[\ell, h]$, then $\hat{\gamma}(v)$ is linear and $\hat{\gamma}'(v) = 0$ in (ℓ, h).

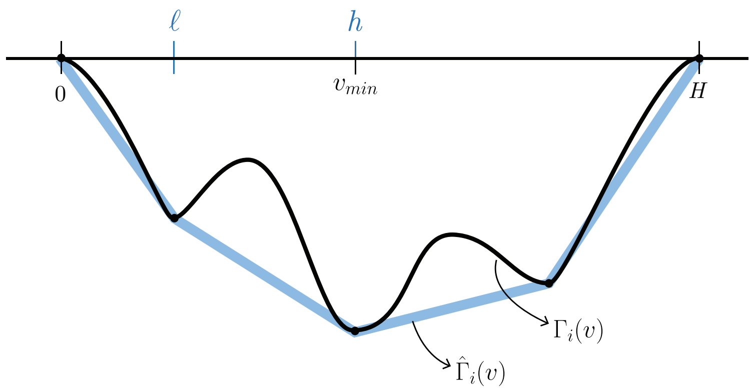


Figure 4: The black curve is $\Gamma_i(v)$, and its lower convex envelope $\hat{\Gamma}(v)$ is traced out by the thick light blue line. The curve is ironed in the interval $[\ell, h]$ (among others), so in that interval, $\hat{\Gamma}(v)$ is linear, and thus has second derivative equal to 0.

We next define the sequence of functions that we will need for the construction:

Definition 3. Let

$$\Gamma_{\geq m}(v) := \Gamma_m(v) \quad \text{and} \quad r_{\geq m} := \text{argmin}_v \Gamma_{\geq m}(v).$$

Inductively, define, for $i := m-1$ down to 1,

$$\Gamma_{\geq i}(v) := \begin{cases} \Gamma_i(v) + \hat{\Gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \Gamma_i(v) + \hat{\Gamma}_{\geq i+1}(r_{\geq i+1}) & v \geq r_{\geq i+1} \end{cases} \quad \text{and} \quad r_{\geq i} := \text{argmin}_v \Gamma_{\geq i}(v).$$

The derivative of $\Gamma_{\geq i}(\cdot)$ is then

$$\gamma_{\geq i}(v) := \begin{cases} \gamma_i(v) + \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \gamma_i(v) & v \geq r_{\geq i+1} \end{cases}.$$

Rewriting this yields

$$\gamma_{\geq i}(v) - \gamma_i(v) = \begin{cases} \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases}. \tag{9}$$

Consider when $m = 2$. Since $\Gamma_i(\cdot)$ is the negative revenue curve for i-day shipping scaled by the probability q_i of drawing a customer with deadline i, then when $m = 2$, $\hat{\Gamma}_2(\cdot)$ is the scaled, negative, ironed, revenue curve for i-day shipping, and we aim to minimize it.

Now, we can observe that $\Gamma_{\geq 1}(v)$ is precisely the revenue from setting a price of v for 1-day shipping and optimizing the revenue for 2-day shipping constrained by the price of v set for 1-day shipping. Under this constraint, the best revenue for 2-day shipping is attained by implementing the optimal price from the ironed revenue curve for 2-day shipping that is at most v (using a lottery if needed). Then deadline 2 customers contribute revenue $-1 \cdot \min_{p_2 \leq v} \hat{\Gamma}_2(p_2)$. Observe the pink curve in Figure 3: since $\hat{\Gamma}_2(\cdot)$ is concave, the minimum is achieved at $\min\{v, r_2\}$. This is exactly what $\Gamma_{\geq 1}(v)$ accounts for: if $v \leq r_2$, then we get the revenue from setting a price of v for 1-day shipping and the ironed revenue of v for 2-day shipping (possibly via lottery), earning $\Gamma_1(v) + \hat{\Gamma}_2(v)$. If $v \geq r_2$, then we get the revenue from setting a price of v for 1-day shipping and from setting the price of r_2 for 2-day shipping.

For the general case, intuitively, these combined curves account for the fact that when i-day shipping's price is low enough to interfere (with respect to the inter-day IC constraint) with the prices that we would like to set for options $i + 1$ through m, we need to consider the problem of setting all of these prices together. However, when i-day's price is high enough not to interfere with the later days, we can just use the optimal choice on days $i + 1$ through m (from $\Gamma_{\geq i+1}(r_{\geq i+1})$) and worry about i-day shipping separately. They also take into account the ironing needed to ensure incentive compatibility.

We can draw an analogy to the ironing in Myerson's optimal auction for irregular distributions. Using the ironed curves ensures incentive compatibility and gives an upper bound on the optimal revenue. Myerson shows that this upper bound is in fact achievable using randomization. Similarly, our combined and ironed curves yield upper bounds on the revenue, and we show how to actually achieve these upper bounds by implementing lotteries.

4.2 The allocation rule

We define the allocation curves $a_i(\cdot)$ inductively. We use the curve $\Gamma_{\geq i}(\cdot)$ and the constraint from the $(i + 1)$-day allocation rule to achieve exactly the revenue that the $\Gamma_{\geq i}(\cdot)$ curves suggest. We will show later that they are optimal. Each allocation curve is piecewise constant. For 1-day shipping, set

$$a_1(v) = \begin{cases} 0 & \text{if } v < r_{\geq 1}, \\ 1 & \text{otherwise.} \end{cases}$$

Suppose that a_{i-1} has been defined for some $i < m$, with jumps at $v_1, \ldots, v_k$, and values $0 = \beta_0 < \beta_1 \leq \beta_2 \ldots \leq \beta_k = 1$. That is,

$$a_{i-1}(v) = \begin{cases} 0 & \text{if } v < v_1, \\ \beta_j & v_j \leq v < v_{j+1} \quad 1 \leq j < k \\ 1 & v_k \leq v. \end{cases}$$

Thus, we can write

$$a_{i-1}(v) = \sum_{j=1}^{k} (\beta_j - \beta_{j-1}) a_{i-1,j}(v)$$

where

$$a_{i-1,j}(v) = \begin{cases} 0 & \text{if } v < v_j \\ 1 & v \geq v_j. \end{cases}$$

Next we define $a_i(v)$.

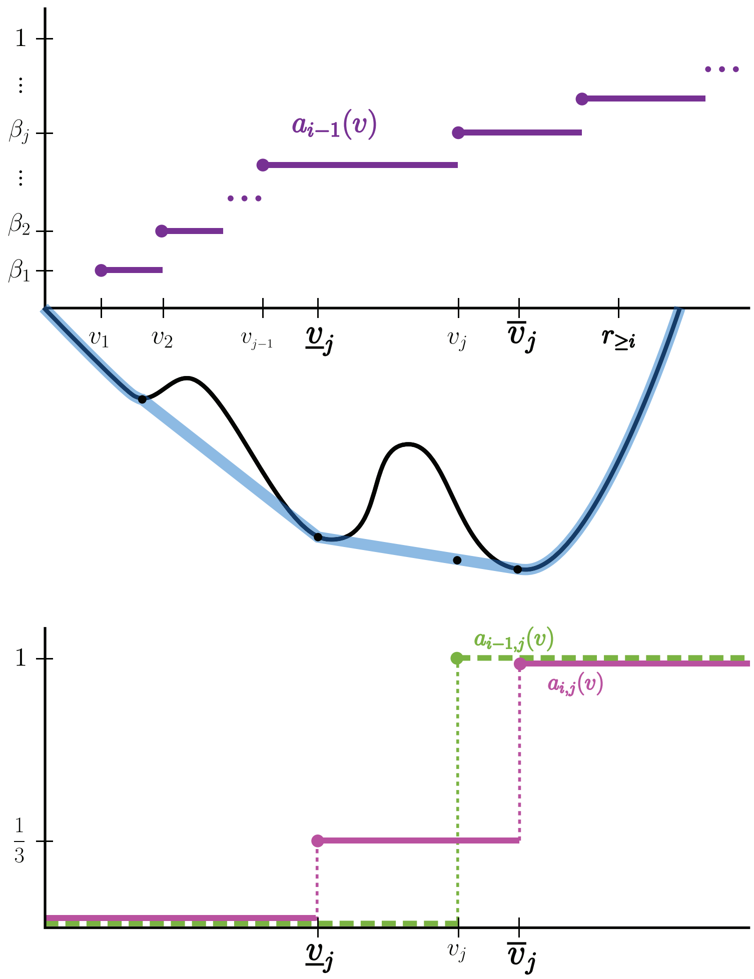


Figure 5: This figure shows an example allocation curve $a_{i-1}(v)$ in purple, and illustrates some aspects of Definition 4. The curves $\Gamma_{\geq i}(v)$ and $\hat{\Gamma}_{\geq i}(v)$ are shown directly below the top figure. In this case, $r_{\geq i} \in [v_{j+1}, v_{j+2})$, so $j^* = j + 1$. The bottom figure shows how $a_{i,j}(v)$ is constructed from $a_{i-1,j}(v)$.

Definition 4. Let j^* be the largest j such that $v_j \leq r_{\geq i}$. For any $j \leq j^*$, consider two cases:

- $\hat{\Gamma}_{\geq i}(v_j) = \Gamma_{\geq i}(v_j)$, i.e. $\hat{\Gamma}_{\geq i}$ not ironed at v_j: In this case, define

$$a_{i,j}(v) = \begin{cases} 0 & \text{if } v < v_j \\ 1 & \textit{otherwise}. \end{cases}.$$

- $\hat{\Gamma}_{\geq i}(v_j) \neq \Gamma_{\geq i}(v_j)$: In this case, let
  - $\underline{v}_j :=$ the largest $v < v_j$ such that $\hat{\Gamma}_{\geq i}(v) = \Gamma_{\geq i}(v)$ i.e., not ironed, and
  - $\overline{v}_j :=$ the smallest $v > v_j$ such that $\hat{\Gamma}_{\geq i}(v) = \Gamma_{\geq i}(v)$ i.e., not ironed.

  Let $0 < \delta < 1$ such that
  $$v_j = \delta \underline{v}_j + (1-\delta)\overline{v}_j.$$
  Then $\hat{\Gamma}_{\geq i}(\cdot)$ is linear between $\underline{v}_j$ and $\overline{v}_j$:
  $$\hat{\Gamma}_{\geq i}(v_j) = \delta \Gamma_{\geq i}(\underline{v}_j) + (1-\delta)\Gamma_{\geq i}(\overline{v}_j).$$
  Define
  $$a_{i,j}(v) = \begin{cases} 0 & \text{if } v < \underline{v}_j \\ \delta & \underline{v}_j \leq v < \overline{v}_j \\ 1 & \textit{otherwise.} \end{cases}$$

Finally, set $a_i(v)$ as follows:

$$a_i(v) = \begin{cases} \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i,j}(v) & \text{if } v < r_{\geq i}, \\ \\ 1 & v \geq r_{\geq i}. \end{cases} \tag{10}$$

Remark: In order to continue the induction and define $a_{i+1}(v)$ we need to rewrite $a_i(v)$ in terms of functions $a_{i,j}(v)$ that take only 0/1 values. This is straightforward.

An Alternate Description: Note that it is equivalent to view $a_{i-1}(\cdot)$ as a randomization over prices where a price of v_j is offered with probability $(\beta_j - \beta_{j-1})$. For each possible price v_j on day $i-1$, we select the optimal choice for i-day shipping using the negative revenue curve $\hat{\Gamma}_{\geq i}(\cdot)$. That is, we determine $p_{i,j} = \text{argmin}_{p \leq v_j} \hat{\Gamma}_{\geq i}(p)$, which is equal to the best constrained price less than v_j to set for i-day shipping to earn revenue for all deadlines i through m, and we implement this price $p_{i,j}$. Note that by convexity, $p_{i,j} = \min\{r_{\geq i}, v_j\}$. In the case that $p_{i,j} = v_j$ and $\hat{\Gamma}_{\geq i}(\cdot)$ is ironed at v_j, this price is implemented by randomizing over $\underline{v}_j$ with probability δ and $\overline{v}_j$ with probability $1-\delta$. We multiply each price (or randomized two prices) for i-day shipping with the probability that v_j was offered for $i-1$-day shipping, giving a randomization over prices for i-day shipping as well, resulting in $a_i(\cdot)$.

In Section 3 we mention that for $m = 2$ days, when the reserve prices are increasing, that is, $r_1 < r_2$, and both distributions F_1 and F_2 are regular, the optimal mechanism sets a single price. We now see that the optimal auction sets a price of $r_{\geq 1}$ on 1-day. Since F_2 is regular, the negative revenue curve $\Gamma_2(\cdot)$ is convex, so there is no ironing. Then the best 2-day option is always a price. Furthermore, because F_1 is regular and thus $\Gamma_1(\cdot)$ is convex with its minimum $r_1 < r_2$, the minimum $r_{\geq 1}$ of $\Gamma_{\geq 1}$ will occur between r_1 and r_2. Since $r_{\geq 1} \leq r_2$, then the best price on 2-day will be precisely $r_{\geq 1}$, so there will be a single price.

Note that for each possible price on day $i-1$ we could offer as many as two prices on i-day, hence there is a trivial upper bound of 2^{i-1} options for i-day shipping. Saxena et al. [2018] provides a matching lower bound.

In the special case that each combined curve $\Gamma_{\geq i}(\cdot)$ is convex for each i, then the allocation rule will be a deterministic pricing. This occurs when each scales revenue curve is concave, or when $f_i(v)\varphi_i(v)$ is monotone weakly increasing. This condition is called *declining marginal revenues* (DMR).[7] However, even in this case, determining the allocation rule is more complicated than one might imagine at first glance. For example, for three deadlines, the optimal deterministic mechanism can require 1, 2, or 3 distinct prices, and determining how many prices to use and how to set them is non-trivial using standard revenue curve approaches. However, our definition of the curves $\Gamma_{\geq i}(\cdot)$ makes determining the optimal pricing immediate.

Lemma 1. *The allocation curves $a_i(\cdot)$, for $1 \leq i \leq m$, are monotone increasing from 0 to 1 and satisfy the inter-day IC constraints (5). Moreover, each $a_i(\cdot)$ changes value only at points where $\hat{\Gamma}_{\geq i}(\cdot)$ is not ironed.*

Proof. That the allocation curves $a_i(\cdot)$ are weakly increasing, start out at 0, and end at 1 is immediate from the fact that they are convex combinations of the monotone allocation curves $a_{i,j}(\cdot)$. Also, by construction, each $a_i(\cdot)$ changes value only at points where $\hat{\Gamma}_{\geq i}(v)$ is not ironed.

So we have only left to verify that

$$\int_0^v a_{i-1}(x)dx \leq \int_0^v a_i(x)dx.$$

From the discussion above, for $v \leq r_{\geq i}$, we have

$$a_{i-1}(v) = \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i-1,j}(v) \quad \text{and} \quad a_i(v) = \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i,j}(v)$$

since $a_{i-1,j}(v) = 0$ for $v \leq r_{\geq i}$ and $j > j^*$. Thus, it suffices to show that for each $j \leq j^*$ and $v \leq r_{\geq i}$

$$\int_0^v a_{i-1,j}(x)dx \leq \int_0^v a_{i,j}(x)dx.$$

If $\hat{\Gamma}_{\geq i}$ is not ironed at v_j, then this is an equality. Otherwise, for $v \leq v_j$, the left hand side is 0 and the right hand side is nonnegative. For $v_j \leq v \leq \overline{v}_j$ the left hand side is $(v - v_j)$, whereas the right hand side is $\delta(v - \underline{v}_j)$. Rearranging the inequality $v_j = \delta\underline{v}_j + (1-\delta)\overline{v}_j \geq \delta\underline{v}_j + (1-\delta)v$ implies that $v - v_j \leq \delta(v - \underline{v}_j)$. This completes the proof that (5) holds.

Notice that $\int_0^v a_{i-1,j}(x)dx = \int_0^v a_{i,j}(x)dx$ for $v < \underline{v}_j$ and $v > \overline{v}_j$, so $a_{i-1}(v) = a_i(v)$ unless $\Gamma_{\geq i}$ is ironed at v, or $v \geq r_{\geq i}$. We will use this fact in the proof of Claim 4 below. □

5 Proof of optimality

In this section, we prove that the allocation rules and pricing of the previous section are optimal. To this end, we formulate our problem as an (infinite) linear program. We discussed the objective and constraints of the primal program in Section 2, and we have already shown above that our allocation rules are feasible for the primal program. We then construct a dual program, and a feasible dual solution for which complementary slackness holds. This implies strong duality holds, and thus, that our solution is optimal.

[7] A one-dimensional distribution F satisfies Declining Marginal Revenues if $v(1?F(v))$ is concave. See Devanur et al. [2017] for examples and more discussion. For an example, uniform distributions are DMR, along with any distribution of bounded support and monotone non-decreasing density.

5.1 The linear programming formulation

Recall the definitions from Section 2: The function $\gamma_i(v)$ is the derivative of $\Gamma_i(v) = \int_0^v q_i\varphi_i(x)f_i(x)\,dx$, where $\varphi_i(v) = v - \frac{1-F_i(v)}{f_i(v)}$ is the i-day virtual value function and q_i is the fraction of bidders with deadline i. Similarly $\hat{\gamma}_i(v)$ is the derivative of $\hat{\Gamma}_i(v)$. We use $[m]$ to denote the set of integers $\{1,\ldots,m\}$.

The Primal

Variables: $a_i(v)$, for all $i \in [m]$, and all $v \in [0, H]$.

$$\text{Maximize } \sum_{i=1}^{m} \int_0^H a_i(v)\gamma_i(v)dv$$

Subject to

$$\begin{aligned}
\int_0^v a_i(x)dx - \int_0^v a_{i+1}(x)dx \le 0 \qquad & \forall i \in [m-1] \quad \forall v \in [0,H] \qquad & \text{(dual variables } \alpha_i(v)) \\
a_i(v) \le 1 \qquad & \forall i \in [m] \quad \forall v \in [0,H] \qquad & \text{(dual variables } b_i(v)) \\
-a_i'(v) \le 0 \qquad & \forall i \in [m] \quad \forall v \in [0,H] \qquad & \text{(dual variables } \lambda_i(v)) \\
a_i(v) \ge 0 \qquad & \forall i \in [m] \quad \forall v \in [0,H]. &
\end{aligned}$$

Note that $a_i'(v)$ denotes $\frac{d}{dv}a_i(v)$.

The Dual

Variables: $b_i(v), \lambda_i(v)$, for all $i \in [m]$, and all $v \in [0, H]$, $\alpha_i(x)$ for $i \in [m-1]$ and all $x \in [0, H]$.

$$\text{Minimize } \int_0^H [b_1(v) + \cdots + b_m(v)]\,dv$$

Subject to

$$\begin{aligned}
b_1(v) + \lambda_1'(v) + \int_v^H \alpha_1(x)dx \ge \gamma_1(v) \qquad & \forall v \in [0,H] \text{ (primal var } a_1(v)) \\
b_i(v) + \lambda_i'(v) + \int_v^H \alpha_i(x)dx - \int_v^H \alpha_{i-1}(x)dx \ge \gamma_i(v) \qquad & \forall v \in [0,H], i = 2,\ldots,m-1 \\
& \text{(primal var } a_i(v)) \\
b_m(v) + \lambda_m'(v) - \int_v^H \alpha_{m-1}(x)dx \ge \gamma_m(v) \qquad & \forall v \in [0,H] \text{ (primal var } a_m(v)) \\
\lambda_i(H) = 0 \qquad & \forall i \in [m] \\
\alpha_i(v) \ge 0 \qquad & \forall v \in [0,H], i \in [m-1] \\
b_i(v), \lambda_i(v) \ge 0 \qquad & \forall i \in [m] \forall v \in [0,H].
\end{aligned}$$

Note that $\lambda_i'(v)$ denotes $\frac{d}{dv}\lambda_i(v)$.

5.2 Conditions for strong duality

As long as there are feasible primal and dual solutions satisfying the following conditions, strong duality holds. See Appendix A for a proof that these conditions are sufficient.

$$a_i(v) > 0 \Rightarrow \lambda_i(v) \text{ continuous at } v \quad i \in [m] \tag{11}$$

$$a_i(v) < 1 \Rightarrow b_i(v) = 0 \quad i \in [m] \tag{12}$$

$$a'_i(v) > 0 \Rightarrow \lambda_i(v) = 0 \quad i \in [m] \tag{13}$$

$$\int_0^v a_i(x)dx < \int_0^v a_{i+1}(x)dx \Rightarrow \alpha_i(v) = 0 \quad i \in [m-1] \tag{14}$$

$$b_i(v) + \lambda'_i(v) + \int_v^H \alpha_i(x)dx - \int_v^H \alpha_{i-1}(x)dx > \gamma_i(v) \Rightarrow a_i(v) = 0 \quad i = 2, \dots, m-1 \tag{15}$$

$$b_1(v) + \lambda'_1(v) + \int_v^H \alpha_1(x)dx > \gamma_1(v) \Rightarrow a_1(v) = 0 \tag{16}$$

$$b_m(v) + \lambda'_m(v) - \int_v^H \alpha_{m-1}(x)dx > \gamma_m(v) \Rightarrow a_m(v) = 0 \tag{17}$$

We allow $a'_i(v) \in \mathbb{R} \cup \{+\infty\}$, otherwise we could not even encode a single-price auction.[8]

5.3 The proof

Theorem 1. *The allocation curves presented in Subsection 4.2 are optimal, that is, obtain the maximum possible expected revenue.*

Proof. To prove the theorem, we verify that there is a setting of feasible dual variables for which all the conditions for strong duality hold. To this end, set the variables as follows:

$$\lambda_i(v) = \Gamma_{\geq i}(v) - \hat{\Gamma}_{\geq i}(v) \tag{18}$$

$$b_i(v) = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v) & v \geq r_{\geq i} \end{cases} \tag{19}$$

$$\alpha_i(v) = \begin{cases} \hat{\gamma}'_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases} \tag{20}$$

The dual variables are selected precisely to satisfy complementary slackness conditions and therefore ensure optimality.

The dual variable $\lambda_i(\cdot)$ corresponds to the monotonicity constraint on $a_i(\cdot)$ in the primal. Since $\Gamma_{\geq i}(\cdot)$ is the curve used to set a_i, it is intuitive the dual variable $\lambda_i(v)$ corresponds to how much we needed to iron $\Gamma_{\geq i}(v)$ for $a_i(\cdot)$ to be monotone at v.

There are m constraints (other than non-negativity) in the dual program, one corresponding to each deadline. We set them so that the constraint for m-day shipping is satisfied with equality.

[8] In particular, $a_i(v)$ may have (countably many) discontinuities, in which points $a'_i(v) = +\infty > 0$. However, in our proof of optimality $a'_i(v)$ appears only as a factor of the product $a'_i(v)\lambda_i(v)$. Every time $a'_i(v) = +\infty$, the corresponding dual value of $\lambda_i(v)$ is 0—by condition (13). See also Appendix A.

We can then add constraints m and $m-1$, and the remaining set of $m-1$ constraints correspond precisely to an $m-1$ deadline problem. Herein lies the basis for the induction.

From Claim 1, it follows that $\lambda_i(v), \alpha_i(v) \geq 0$ for all v and i. Since $r_{\geq i}$ is the minimum of $\hat{\Gamma}_{\geq i}(\cdot)$, we have $\hat{\gamma}_{\geq i}(r_{\geq i}) = 0$. Moreover, since $\hat{\gamma}_{\geq i}(\cdot)$ is increasing, $b_i(v) \geq 0$ for all v and i.

Taking the derivative of (18), and using Equation (9), we obtain:

$$\gamma_i(v) - \lambda_i'(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) - \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \hat{\gamma}_{\geq i}(v) - 0 & v \geq r_{\geq i+1} \end{cases} \tag{21}$$

$$\gamma_m(v) - \lambda_m'(v) = \hat{\gamma}_m(v) \tag{22}$$

Also, using (20) and the fact that $\hat{\gamma}_{\geq i+1}(r_{\geq i+1}) = 0$, we get:

$$A_i(v) := \int_v^H \alpha_i(x)\, dx = \begin{cases} -\hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases} \tag{23}$$

Condition (11) from Section 5.2 holds since $\Gamma_{\geq i}(v)$ and $\hat{\Gamma}_{\geq i}(v)$ are both continuous functions. The proofs of all remaining conditions for strong duality from Section 5.2 can be found below. □

Claim 2. Condition (12): For all i and v, $a_i(v) < 1 \implies b_i(v) = 0$.

Proof. If $a_i(v) < 1$, then $v < r_{\geq i}$, so by construction, $b_i(v) = 0$. □

Claim 3. Condition (13): For all i and v, $a_i'(v) > 0 \implies \lambda_i(v) = 0$.

Proof. From Subsection 4.2, $a_i'(v) > 0$ only for unironed values of v, at which $\lambda_i(v) = 0$. □

Claim 4. Condition (14): For all i and v, $\int_0^v a_i(x)dx < \int_0^v a_{i+1}(x)dx \implies \alpha_i(v) = 0$.

Proof. As discussed at the end of the proof of Lemma 1, $\int_0^v a_i(x)dx = \int_0^v a_{i+1}(x)dx$ unless $\Gamma_{\geq i+1}$ is ironed at v, or $v \geq r_{\geq i}$. In both of these cases $\alpha_i(v) = 0$ (by part 4 of Claim 1 and Definition 20, respectively). □

Claim 5. Conditions (15)- (17) and dual feasibility: For all i and v, $a_i(v) > 0 \implies$ the corresponding dual constraint is tight, and the dual constraints are always feasible.

Proof. Rearrange the dual constraint $b_i(v) + A_i(v) - A_{i-1}(v) + \lambda_i'(v) \geq \gamma_i(v)$ to

$$b_i(v) - A_{i-1}(v) \geq \gamma_i(v) - \lambda_i'(v) - A_i(v).$$

Fact 1: For $i \in [m-1]$, $\gamma_i(v) - \lambda_i'(v) - A_i(v) = \hat{\gamma}_{\geq i}(v)$ for all v. To see this use (21) and (23):

$$\gamma_i(v) - \lambda_i'(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) - \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \hat{\gamma}_{\geq i}(v) - 0 & v \geq r_{\geq i+1} \end{cases} \qquad A_i(v) = \begin{cases} -\hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases}$$

Fact 2: For $i \in \{2, \ldots, m\}$, $b_i(v) - A_{i-1}(v) = \hat{\gamma}_{\geq i}(v)$ for all v.

$$b_i(v) = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v) & v \geq r_{\geq i} \end{cases} \qquad -A_{i-1}(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) & v < r_{\geq i} \\ 0 & v \geq r_{\geq i} \end{cases}$$

Hence for $i \in \{2, \ldots, m-1\}$, $b_i(v) - A_{i-1}(v) = \gamma_i(v) - \lambda'_i(v) - A_i(v)$ for all v.

For $i = m$, since $\gamma_{\geq m} = \gamma_m$, and $\gamma_m(v) - \lambda'_m(v) = \hat{\gamma}_m(v)$. Combining this with Fact 2 above, we get that $b_m(v) - A_{m-1}(v) + \lambda'_m(v) = \gamma_m(v)$ for all v.

Finally, for $i = 1$, using Fact 1, for $v < r_{\geq 1}$, we get

$$b_1(v) = 0 \geq \hat{\gamma}_{\geq 1}(v) = \gamma_1(v) - \lambda'_1(v) - A_1(v)$$

which is true for $v < r_{\geq 1}$. For $v \geq r_{\geq 1}$, we get

$$b_1(v) = \gamma_{\geq 1}(v) = \gamma_1(v) - \lambda'_1(v) - A_1(v),$$

so the dual constraint is tight when $a_1(v) > 0$ as this starts at $r_{\geq 1}$. □

The above claims prove that this dual solution satisfies feasibility and all complementary slackness and continuity conditions from Section 5.2 hold.

6 Closed-Form Virtual Values

By taking the partial Lagrangian of our primal from subsection 5.1, we can view the optimal mechanism as an expected virtual welfare maximizer. By plugging in our closed-form dual variables, we produce closed-form virtual value functions. We use our above approach combined with the approximation of Cai et al. [2016], only used for optimal revenue instead of approximation.

We multiply each constraint (aside from feasibility) by its dual variable and move it into the objective function, minimizing over these non-negative dual variables. The basics of Lagrangian duality are outlined in Appendix B. The resulting partial Lagrangian primal is:

$$\max_{a_i(v) \in [0,1]} \quad \min_{\lambda_i(v), \alpha_i(v) \geq 0} \mathcal{L}(a, \lambda)$$

where

$$\begin{aligned}\mathcal{L}(a, \lambda) := \sum_{i=1}^{m} \int_0^H a_i(v)\gamma_i(v)dv + \sum_{i=1}^{m-1} \int_0^v \alpha_i(v) \left[\int_0^v a_{i+1}(x)dx - \int_0^v a_i(x)dx\right] dv \\ + \sum_{i=1}^{m} \int_0^v \lambda_i(v) \left[a'_i(v)\right] dv\end{aligned}$$

Recall that $A_i(v) = \int_v^H \alpha_i(x)dx$. Then, using this notation, as well as integration by parts on the λ terms, and aggregating the a terms, we can rewrite $\mathcal{L}$ as follows. (This is similar to the steps

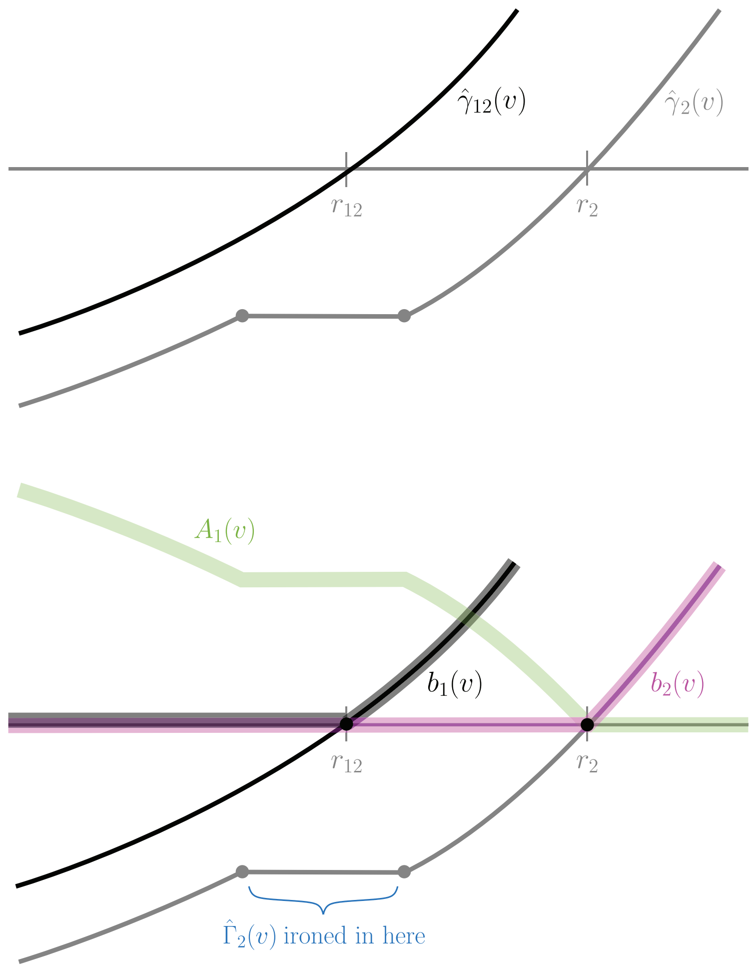


Figure 6: This figure illustrates what some of the dual variables might be for the case of two days when $r_{\geq 1} < r_2$. The upper figure plots the functions $\hat{\gamma}_{12}(v)$ and $\hat{\gamma}_2(v)$, and the lower figure shows $b_1(v)$ in dark grey, $b_2(v)$ in pink and $A_1(v) = \int_v^H \alpha_1(x)dx$ in green. Note that up to r_2, the function $A_1(v) = -\hat{\gamma}_2(v)$.

we take in the proof of strong duality in Appendix A.)

$$\begin{aligned}
\mathcal{L}(a,\lambda) = \int_0^H & \left[a_1(v)\gamma_1(v) + \lambda_1(v)a_1'(v) - a_1(v)A_1(v)\right] dv \\
& + \sum_{i=2}^{m-1} \int_0^H \left[a_i(v)\gamma_i(v) + \lambda_i(v)a_i'(v) + a_i(v)A_{i-1}(v) - a_i(v)A_i(v)\right] dv \\
& + \int_0^H \left[a_m(v)\gamma_m(v) + \lambda_m(v)a_m'(v) + a_m(v)A_{m-1}(v)\right] dv \\
= \int_0^H & f_1(v)a_1(v) \left[\frac{\gamma_1'(v)}{f_1(v)} - \frac{\lambda_1'(v)}{f_1(v)} - \frac{A_1(v)}{f_1(v)}\right] dv \\
& + \sum_{i=2}^{m-1} \int_0^H f_i(v)a_i(v) \left[\frac{\gamma_i'(v)}{f_i(v)} - \frac{\lambda_i'(v)}{f_i(v)} - \frac{A_i(v)}{f_i(v)} + \frac{A_{i-1}(v)}{f_i(v)}\right] dv \\
& + \int_0^H f_m(v)a_m(v) \left[\frac{\gamma_m'(v)}{f_m(v)} - \frac{\lambda_m'(v)}{f_m(v)} + \frac{A_{m-1}(v)}{f_m(v)}\right] dv
\end{aligned}$$

This gives that $\mathcal{L}$ is equal to expected virtual welfare for the following virtual values:

$$\phi_i(v) = \frac{\gamma_i'(v)}{f_i(v)} - \frac{\lambda_i'(v)}{f_i(v)} - \frac{A_i(v)}{f_i(v)} + \frac{A_{i-1}(v)}{f_i(v)} \qquad \text{for } i \in \{2, \ldots, m-1\},$$

$$\phi_1(v) = \frac{\gamma_1'(v)}{f_1(v)} - \frac{\lambda_1'(v)}{f_1(v)} - \frac{A_1(v)}{f_1(v)}, \qquad \text{and} \qquad \phi_m(v) = \frac{\gamma_m'(v)}{f_m(v)} - \frac{\lambda_m'(v)}{f_m(v)} + \frac{A_{m-1}(v)}{f_m(v)}.$$

Recall that our partial Lagrangian primal is of the form

$$\max_{a_i(v)\in[0,1]} \quad \min_{\lambda_i(v),\alpha_i(v)\geq 0} \mathcal{L}(a,\lambda) = \max_{a_i(v)\in[0,1]} \quad \min_{\lambda_i(v),\alpha_i(v)\geq 0} \sum_{i=1}^{m} f_i(v)a_i(v)\phi_i(v).$$

Note that ϕ depends on λ, α, and we must find the variables that minimize these functions. We plug in our optimal dual variables from our closed-form solution to FedEx, giving closed-form virtual values:

$$\phi_1 = \hat{\gamma}_1(v)/f_1(v) \qquad \text{and} \qquad \phi_i = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v)/f_i(v) & v \geq r_{\geq i} \end{cases} \qquad \text{for } i \in \{2, \ldots, m\}.$$

Then, the allocation rule that maximizes expected virtual welfare for these virtual value functions is precisely the optimal mechanism. Note that because our closed-form solution exists, strong duality holds, so the Lagrangian primal is not relaxed.

7 Interdimensional Settings

Many other natural problems fall into this category of "interdimensional" as well. Consider a buyer with a value for an item and a private budget b which is the most that he can pay [Devanur and Weinberg, 2017]. Or, suppose a buyer has a value v for each unit of an item up to some private demand capacity d [Devanur et al., 2017].

We highlight now some of the features of the FedEx setting that are common also to the single-minded setting in [Devanur et al., 2019], as well as these other interdimensional settings. These properties also help to explain why duality techniques gain traction.

- Every allocation rule (which lists, for each (value, deadline) pair, a probability of receiving each of the three items) can be "collapsed" to simply list, for each (value, deadline) pair, a single probability (of receiving a satisfying item).
- Local Incentive Compatibility (IC) constraints imply global IC constraints. That is, any auction satisfying all local IC constraints is also globally IC.
- A payment identity applies: a simple closed form determines payments as a function of the allocation rule.

Of course, these three properties are intertwined: without a collapsible allocation rule, no closed-form payment identity is possible.

There are other commonalities as well. In each of these settings, it has been shown (in this paper, Devanur et al. [2019], [DW '17], and [DHP '17]) that the optimal mechanism is deterministic when the marginal distributions satisfy declining marginal revenues (DMR). These works also show that the degree of randomization (the menu complexity) of the optimal mechanism is larger than that in single-dimensional settings, but smaller than in multi-dimensional settings, for the FedEx, budgets, and the single-minded settings.

References

Alaei, S. (2011). Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 512–521.

Anderson, E. J. and Nash, P. (1987). *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley & Sons.

Armstrong, M. (1996). Multiproduct Nonlinear Pricing. *Econometrica*, 64(1):51–75.

Babaioff, M., Immorlica, N., Lucier, B., and Weinberg, S. M. (2014). A Simple and Approximately Optimal Mechanism for an Additive Buyer. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 21–30. IEEE.

Bergemann, D., Brooks, B., and Morris, S. (2015). The Limits of Price Discrimination . *American Economic Review*, 105(3):921–957.

Briest, P., Chawla, S., Kleinberg, R., and Weinberg, S. M. (2015). Pricing lotteries. *Journal of Economic Theory*, 156:144–174.

Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *Journal of political economy*, 97(5):1060–1090.

Cai, Y. and Daskalakis, C. (2011). Extreme-value theorems for optimal multidimensional pricing. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 522–531.

Cai, Y., Daskalakis, C., and Weinberg, S. M. (2011). On optimal multidimensional mechanism design. *SIGecom Exch.*, 10(2):29–33.

Cai, Y., Daskalakis, C., and Weinberg, S. M. (2012). Optimal multi-dimensional mechanism design: Reducing revenue to welfare maximization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 130–139.

Cai, Y., Daskalakis, C., and Weinberg, S. M. (2013a). Reducing revenue to welfare maximization: Approximation algorithms and other generalizations. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 578–595. SIAM.

Cai, Y., Daskalakis, C., and Weinberg, S. M. (2013b). Understanding incentives: Mechanism design becomes algorithm design. *CoRR*, abs/1305.4002.

Cai, Y., Devanur, N. R., and Weinberg, S. M. (2016). A Duality Based Unified Approach to Bayesian Mechanism Design. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 926–939, New York, NY, USA. ACM.

Cai, Y. and Huang, Z. (2013). Simple and nearly optimal multi-item auctions. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 564–577.

Chawla, S., Hartline, J. D., and Kleinberg, R. (2007). Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, EC '07, pages 243–251.

Chawla, S., Hartline, J. D., Malec, D. L., and Sivan, B. (2010). Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320. ACM.

Chawla, S., Malec, D., and Sivan, B. (2015). The power of randomness in bayesian optimal mechanism design. *Games and Economic Behavior*, 91:297 – 317.

Chawla, S. and Sivan, B. (2014). Bayesian algorithmic mechanism design. *SIGecom Exch.*, 13(1):5–49.

Daskalakis, C. (2015). Multi-item auctions defying intuition? *ACM SIGecom Exchanges*, 14(1):41–75.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2013). Mechanism design via optimal transport. In *Proceedings of the 14th ACM Conference on Electronic Commerce*, EC '13, pages 269–286.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2017). Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767.

Devanur, N. R., Goldner, K., Saxena, R. R., Schvartzman, A., and Weinberg, S. M. (2019). Optimal mechanism design for single-minded agents. In *Manuscript*.

Devanur, N. R., Haghpanah, N., and Psomas, C.-A. (2017). Optimal Multi-Unit Mechanisms with Private Demands. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 41–42, New York, NY, USA. ACM.

Devanur, N. R. and Weinberg, S. M. (2017). The Optimal Mechanism for Selling to a Budget Constrained Buyer: The General Case. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 39–40, New York, NY, USA. ACM.

Giannakopoulos, Y. and Koutsoupias, E. (2014). Duality and Optimality of Auctions for Uniform Distributions. In *Proceedings of the 15th ACM Conference on Economics and Computation*, EC '14, pages 259–276.

Giannakopoulos, Y. and Koutsoupias, E. (2015). Selling Two Goods Optimally. In *ICALP (2)*, volume 9135 of *Lecture Notes in Computer Science*, pages 650–662. Springer.

Haghpanah, N. and Hartline, J. (2015). Reverse mechanism design. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 757–758. ACM. Updated version: http://arxiv.org/abs/1404.1341.

Hart, S. and Nisan, N. (2013). The menu-size complexity of auctions. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 565–566, New York, NY, USA. ACM.

Hart, S. and Nisan, N. (2017). Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347.

Hart, S. and Reny, P. J. (2012). Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Tech Report*.

Hartline, J. D. (2013). Mechanism design and approximation. *Book draft. October*, 122.

Laffont, J.-J., Maskin, E., and Rochet, J.-C. (1987). Optimal nonlinear pricing with two-dimensional characteristics. *Information, Incentives and Economic Mechanisms*, pages 256–266.

Li, X. and Yao, A. C.-C. (2013). On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237.

Luenberger, D. G. (1997). *Optimization by Vector Space Methods*. John Wiley & Sons.

Manelli, A. M. and Vincent, D. R. (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1 – 35.

McAfee, R. P. and McMillan, J. (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335 – 354.

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

Pavlov, G. (2011). Optimal mechanism for selling two goods. *The B.E. Journal of Theoretical Economics*, 11(3).

Pycia, M. (2006). *Stochastic vs deterministic mechanisms in multidimensional screening*. PhD thesis, MIT. Chapter 3.

Riley, J. and Zeckhauser, R. (1983). Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289.

Rockafellar, R. T. (1974). *Conjugate duality and optimization*, volume 16. Siam.

Roughgarden, T. (2015). Approximately optimal mechanism design: Motivation, examples, and lessons learned. *SIGecom Exch.*, 13(2):4–20.

Saxena, R. R., Schvartzman, A., and Weinberg, S. M. (2018). The menu-complexity of one-and-a-half dimensional mechanism design. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18.

Thanassoulis, J. (2004). Haggling over substitutes. *Journal of Economic Theory*, 117(2):217 – 245.

Yao, A. C.-C. (2015). An n-to-1 Bidder Reduction for Multi-item Auctions and its Applications. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 92–109. SIAM.

A Proof of strong duality

Theorem 2. *Let $a_i(\cdot)$, $b_i(\cdot)$, $\lambda_i(\cdot)$, $\alpha_i(\cdot)$ be functions feasible for the primal and dual, satisfying all the conditions from Section sec:CS. Then they are optimal.*

Proof. First, we prove weak duality. For any feasible primal and dual:

$$\int_0^H \sum_{i=1}^m b_i(v)\,dv \tag{24}$$

$$= \int_0^H \sum_{i=1}^m (1 \cdot b_i(v) + 0 \cdot [\lambda_i(v) + \alpha_i(v)])\ dv. \tag{25}$$

Applying primal feasibility, we see that this quantity is

$$\geq \int_0^H \sum_{i=1}^m \left(a_i(v) b_i(v)) - a_i'(v)\lambda_i(v) + \left[\int_0^v a_i(x) - a_{i+1}(x)dx\right] \alpha_i(v) \right)\ dv. \tag{26}$$

We rewrite this expression using the following.

- Applying integration by parts,using the facts that $\lambda_i(\cdot)$ is continuous (Condition (11)) and equal to 0 at any point that $a_i'(v) = \infty$,[9] we get
$$-\int_0^H a_i'(v)\lambda_i(v)\,dv = -a_i(v)\lambda_i(v)\Big|_0^H + \int_0^H a_i(v)\lambda_i'(v)\,dv = \int_0^H a_i(v)\lambda_i'(v)\,dv,$$
since $a_i(0) = 0$ and $\lambda_i(H) = 0$.

- Second, interchanging the order of integration, we get
$$\int_0^H \int_0^v [a_i(x) - a_{i+1}(x)dx]\,\alpha_i(v)\,dv = \int_0^H \left(a_i(v) \int_v^H \alpha_i(x)\,dx - a_{i+1}(v) \int_v^H \alpha_i(x)\,dx \right)\ dv.$$

Combining these shows that (26) equals

$$\int_0^H \left(\sum_{i=1}^m a_i(v) \left[b_i(v) + \lambda_i'(v) + \int_v^H \alpha_i(x) - \int_v^H \alpha_{i-1}(x)\,dx \right] \right)\ dv$$

$$\geq \int_0^H \sum_{i=1}^H a_i(v)\gamma_i(v)\,dv \tag{27}$$

where the last inequality is dual feasibility. (Note that $\alpha_0(\cdot) = \alpha_m(\cdot) = 0$.)

Comparing (24) and (27) yields weak duality, i.e., $\sum_i \int_0^H b_i(v)\,dv \geq \sum_i \int_0^H a_i(v)\gamma_i(v)\,dv$.

If the conditions (11)-(17) hold, we also have strong duality and hence optimality: To show that (25) = (26), observe that

- (12) $a_i(v) < 1$ implies that $b_i(v) = 0$;
- (13) $a_i'(v) > 0$ implies that $\lambda_i(v) = 0$.
- (14) $\int_0^v (a_{i+1}(x) - a_i(x))dx > 0$ implies that $\alpha_i(v) = 0$ for $i = 1, \ldots, n-1$.

Finally, (27) is an equality rather than an inequality because of conditions (15)-(17). □

[9] $a_i'(v)$ can be ∞ at only countably many points.

B Lagrangian Duality

In this section, we provide the basics surrounding formulating a partial Lagrangian primal and taking its dual, and understanding the properties of duality.

We begin with a standard maximization problem subject to constraints, which we call the full primal. The set $\mathcal{P}$ here denotes feasibility constraints, while x represents whatever our primal variables are.

Full primal:

$$\begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & Ax \leq b && \text{(dual variable } \lambda) \\ & x \in \mathcal{P} \end{aligned}$$

We denote the optimal solution to the full primal as x^*; that is, $x^* \in \text{argmax}_{Ax \leq b, x \in \mathcal{P}} f(x)$.

We now form the partial Lagrangian primal by using the Lagrangian multiplier λ_i for each constraint of the form $(Ax)_i \leq b_i$ and moving it into the objective, where we now minimize over the multipliers λ. We leave all of the feasibility constraints as is, and define $\mathcal{L}(x; \lambda)$ as this new objective.

Lagrangian Primal:

$$\max_{x \in \mathcal{P}} \min_{\lambda \geq 0} \quad \mathcal{L}(x; \lambda) \quad = \quad \max_{x \in \mathcal{P}} \min_{\lambda \geq 0} \quad f(x) + \lambda^T(b - Ax)$$

By reversing the order of the max and the min, we obtain the dual minimization problem. We notate this dual problem as $D(\lambda)$.

Lagrangian Dual:

$$\min_{\lambda \geq 0} \quad D(\lambda) \quad = \quad \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \quad f(x) + \lambda^T(b - Ax)$$

We denote the optimal dual solution as $\lambda^* \in \text{argmin}_{\lambda \geq 0} D(\lambda)$.

We say that x, λ satisfy complementary slackness if $\lambda_i \geq 0 \implies b_i - (Ax)_i = 0$.

Relaxation. First, we observe that the (partial) Lagrangian Primal is in fact a relaxation of the full primal. For any feasible x, λ—that is, $Ax \leq b, x \in \mathcal{P}$, and $\lambda \geq 0$—then $f(x) \leq \mathcal{L}(x; \lambda)$.

Weak Duality. The value of the full primal is always upper-bounded by the value of the dual problem. Specifically, the value of the full primal is at most $f(x^*)$ by definition, and any feasible dual solution must satisfy $\lambda \geq 0$, so the dual objective is larger: $f(x^*) \leq D(\lambda)$.

Proof.

$$\begin{aligned} f(x^*) &\leq f(x^*) + \lambda^T(b - Ax^*) && \lambda \geq 0, Ax^* \leq b \\ &\leq \max_{x \in \mathcal{P}} f(x) + \lambda^T(b - Ax) && x^* \in \mathcal{P} \\ &= D(\lambda) \end{aligned}$$

□

Strong Duality. Strong duality implies that the value of the full primal is equal to the value of the Lagrangian primal, and this is equal to the value of the Lagrangrian dual, when they are all at their optimal solutions. However, strong duality is not a given. We see below that if strong duality holds, there must exist a pair of primal, dual solutions that are optimal. Further, if there exist an optimal pair, then strong duality must hold. Either condition is sufficient to show the other exists.

An Optimal Pair implies Strong Duality. For any choice of dual variables $\hat{\lambda}$, if there exists $\hat{x}$ that forms an optimal pair with $\hat{\lambda}$, that is, $\hat{x}$ such that:

1. $\hat{x} \in \text{argmax}_{x\in\mathcal{P}}\mathcal{L}(x;\hat{\lambda})$ ($\hat{x}$ is optimal)
2. $A\hat{x} \le b$ ($\hat{x}$ satisfies the Lagrangified constraints)
3. $\hat{x}, \hat{\lambda}$ satisfy complementary slackness

then strong duality holds, that is, $D(\hat{\lambda}) = f(x^*)$.

Proof.

$$\begin{aligned} D(\hat{\lambda}) &= \max_{x\in\mathcal{P}} \mathcal{L}(x,\hat{\lambda}) \\ &= f(\hat{x}) + \hat{\lambda}^*(b - A\hat{x}) && \text{by (1)} \\ &= f(\hat{x}) && \text{by (3)} \\ &\le f(x^*) && \text{by (2)}, x \in \mathcal{P} \end{aligned}$$

□

Strong Duality implies an Optimal Pair. If strong duality holds, that is, $\min_{\lambda\ge 0} D(\lambda) = f(x^*)$, then there exists $\hat{x}$ such that

1. $\hat{x} \in \text{argmax}_x\mathcal{L}(x;\lambda^*)$
2. $A\hat{x} \le b$
3. $\hat{x}, \lambda^*$ satisfy complementary slackness
4. $f(\hat{x}) = f(x^*)$.

Proof. From weak duality, we know that

$$\min_{\lambda\ge 0} D(\lambda) \quad = \quad D(\lambda^*) \quad \ge \quad \mathcal{L}(x^*,\lambda^*) \quad \ge \quad f(x^*).$$

These inequalities must all hold with equality for the premise to hold. The first inequality's tightness implies condition (1), and the second inequality's tightness implies condition (3). Condition (2) is true by the definition of x^*. □

For further background on Lagrangian duality, see [Rockafellar, 1974].