Efficient Algorithms for Learning Revenue-Maximizing Two-Part Tariffs

Maria-Florina Balcan[1] , **Siddharth Prasad**[1] and **Tuomas Sandholm**[1,2,3,4]

[1]School of Computer Science, Carnegie Mellon University
[2]Optimized Markets, Inc.
[3]Strategy Robot, Inc.
[4]Strategic Machine, Inc.
{ninamf, sprasad2, sandholm}@cs.cmu.edu

Abstract

A two-part tariff is a pricing scheme that consists of an up-front lump sum fee and a per unit fee. Various products in the real world are sold via a menu, or list, of two-part tariffs—for example gym memberships, cell phone data plans, etc. We study learning high-revenue menus of two-part tariffs from buyer valuation data, in the setting where the mechanism designer has access to samples from the distribution over buyers' values rather than an explicit description thereof. Our algorithms have clear direct uses, and provide the missing piece for the recent generalization theory of two-part tariffs. We present a polynomial time algorithm for optimizing one two-part tariff. We also present an algorithm for optimizing a length-L menu of two-part tariffs with run time exponential in L but polynomial in all other problem parameters. We then generalize the problem to multiple markets. We prove how many samples suffice to guarantee that a two-part tariff scheme that is feasible on the samples is also feasible on a new problem instance with high probability. We then show that computing revenue-maximizing feasible prices is hard even for buyers with additive valuations. Then, for buyers with identical valuation distributions, we present a condition that is sufficient for the two-part tariff scheme from the unsegmented setting to be optimal for the market-segmented setting. Finally, we prove a generalization result that states how many samples suffice so that we can compute the unsegmented solution on the samples and still be guaranteed that we get a near-optimal solution for the market-segmented setting with high probability.

1 Introduction

A *two-part tariff (TPT)* consists of an up-front lump sum fee p_1 and a fee p_2 for every additional unit purchased. Various goods and services are priced using such a scheme. For example, Keurig sells coffee machines (the up-front fee) that require proprietary coffee pods (the per unit fee). Another example is health club memberships, where participants often are required to pay an up-front fixed membership fee, as well as a monthly fee. More generally, a length L menu of TPTs is a list $((p_1^1, p_2^1), \ldots, (p_1^L, p_2^L))$ of L TPTs, and a buyer may elect to pay according to any one of the L TPTs (or not to buy anything). Menus of TPTs are also prevalent: health clubs, amusement parks, wholesale stores like Costco, cell phone companies, and credit card companies all frequently offer various tiers of membership usually consisting of lower future payments for a larger up-front payment.

In an early analysis of TPTs, Oi [1971] inspires the problem via Disneyland trying to decide between charging attendees a hefty entrance fee and allowing them free access to rides, versus charging a nominal entrance fee but requiring payment for each ride. An even earlier discussion of TPTs is given by Lewis [1941], where the merits and drawbacks of TPTs are discussed in contexts such as the telephone system, gas legislation, and the UK Central Electricity Board.

We study the problem of learning high-revenue menus of TPTs from buyer valuation data. This can be viewed as a form of *automated mechanism design* [Conitzer and Sandholm, 2002]. In our setting, the seller has access to samples from the distribution over buyers' values, but not an explicit description thereof. This differs from the usual approach taken by the economic theory literature, and instead takes the *sample-based approach to mechanism design*, introduced by Sandholm and Likhodedov [2004; 2005; 2015]. Balcan, Sandholm, and Vitercik [2018] study the sample complexity of revenue maximization, deriving a broad characterization of the number of samples needed to ensure with high probability that a mechanism that achieves high empirical revenue on the samples also generalizes well, that is, achieves high expected revenue over a freshly drawn sample. Our main goal is to provide efficient algorithms for finding menus of TPTs that achieve high empirical revenue over a given set of samples. Many of the mechanism settings studied by Balcan et al. have large parameter spaces and require a number of samples that is exponential in the problem parameters to guarantee generalization. However, they show that the sample complexity of TPTs has only a mild (at most linear) dependence on the parameters, so it is reasonable to ask for sample efficient and computationally efficient algorithms for finding nearly optimal solutions. We present such algorithms, thereby providing the missing, complementary piece to the results of Balcan et al. Our algorithms also have the obvious practical uses in designing TPTs and menus thereof.

1.1 Problem Formulation

In our model, the seller has K units of a good to sell among n buyers $j \in \{1, \ldots, n\}$ via a menu of TPTs. Each buyer is described by his valuation function $v_j : \{1, \ldots, K\} \to \mathbb{R}$ over the K units. So, $v_j(q)$ is the value that buyer j assigns to getting q units of the item. (We implicitly assume that each buyer's value for getting nothing is zero.) We assume that buyers act in a utility maximizing manner: when presented with a menu $((p_1^1, p_2^1), \ldots, (p_1^L, p_2^L))$ of TPTs, buyer j with valuation function $v_j : \{1, \ldots, K\} \to \mathbb{R}$ will choose to buy q units priced by tariff r to maximize $v_j(q) - (p_1^r + q \cdot p_2^r)$, buying 0 units if there are no values of q and r that make the above expression non-negative. Given one sample of buyers $v = (v_1, \ldots, v_n)$, when faced with menu $\mathbf{p}$, say buyer j purchases quantity q_j of tariff r_j. Then the revenue of $\mathbf{p}$ with respect to v, denoted $\mathrm{Rev}_v(\mathbf{p})$, is $\sum_{j=1}^n \mathbf{1}(q_j \geq 1) \cdot (p_1^{r_j} + q_j \cdot p_2^{r_j})$.[1]

However, the model above allows for the possibility that the total quantity $\sum_{j=1}^n q_j$ is larger than K. To deal with this issue, we will usually stipulate that the seller offers a menu $\mathbf{p}$ that is *feasible* (that is, the total quantity purchased is at most K) for each sample he sees (and we show that doing so ensures with high probability that the menu $\mathbf{p}$ is also feasible on a freshly drawn potential future sample).

We also study the case where each buyer belongs to one of M markets, in which case a TPT pricing scheme is of the form $(\mathbf{p}_1, \ldots, \mathbf{p}_M)$, where buyers in market m are offered menu $\mathbf{p}_m$. Revenue is defined similarly, which we denote by $\mathrm{Rev}_v(\mathbf{p}_1, \ldots, \mathbf{p}_M)$. For a set of samples $S = \{v^1, \ldots, v^N\}$, the empirical revenue of $\mathbf{p}$ with respect to S is denoted by $\widehat{\mathrm{Rev}}_S(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^N \mathrm{Rev}_{v^i}(\mathbf{p})$, and similarly for the market-segmented case.

We now state the formal generalization guarantee of Balcan, Sandholm, and Vitercik [2018] for the mechanism class of length-L menus of TPTs for selling K units to n buyers partitioned into M markets. Let $\mathcal{D}$ be some unknown distribution over n-tuples of buyer valuations and markets. For any $0 < \varepsilon, \delta < 1$, there exists an $N_{TPT}(\varepsilon, \delta) \in \mathbb{N}$ such that for all $N \geq N_{TPT}(\varepsilon, \delta)$, it holds with probability at least $1 - \delta$ over the draw of $S = \{v^1, \ldots, v^N\} \sim \mathcal{D}^N$ that for every M-tuple $(\mathbf{p}_1, \ldots, \mathbf{p}_M)$ of length L menus of TPTs,

$$|\widehat{\mathrm{Rev}}_S(\mathbf{p}_1, \ldots, \mathbf{p}_M) - \mathrm{E}_{v \sim \mathcal{D}}[\mathrm{Rev}_v(\mathbf{p}_1, \ldots, \mathbf{p}_m)| \leq \varepsilon.$$

The sample complexity $N_{TPT}(\varepsilon, \delta)$ is at most $O_{\varepsilon, \delta}(ML \log(nKL))$, where we have hidden the dependence on ε and δ as is typical in learning theory. This follows from the piecewise structure of the class of revenue functions: there is a partition of the TPT parameter space $\mathbb{R}^{2LM}$ by hyperplanes into not-too-numerous regions such that empirical revenue is linear over each region (this notion is formalized in the main result of Balcan et al. [2018]).

The overarching goal of our paper is to efficiently find TPT pricing schemes that maximize empirical revenue over a set of samples—which by the above uniform convergence result is highly likely to be nearly optimal in terms of expected revenue as well. The number of samples needed to guarantee

---

[1] We use boldface $\mathbf{p} \in \mathbb{R}^{2L}$ to abbreviate a menu of L TPTs. It is understood, then, that p_1^r and p_2^r denote $\mathbf{p}_{2r-1}$ and $\mathbf{p}_{2r}$, respectively.

generalization only depends at most linearly in the problem parameters, so computationally efficient algorithms for empirical revenue maximization in this setting will be sample efficient as well.

1.2 Summary of the Contributions of this Paper

In Section 2 we give efficient algorithms for finding the empirical revenue maximizing menu of TPTs when the menu length is a fixed constant. Our main result here is an $O(N^3 K^3)$ algorithm when $L = 1$ in the single buyer setting, that generalizes to an $O(n^3 N^3 K^3)$ algorithm in the multi-buyer setting (Section 2.1). We then give an $(NK)^{O(L)}$ algorithm for the setting where $L \geq 1$ (Section 2.2). This algorithm exploits the geometric structure of the problem—buyers' valuations partition the parameter space into several convex polytopes, and revenue maximization over each polytope reduces to solving a linear program.

In Section 3 we generalize the problem to multiple markets. We prove how many samples suffice to guarantee that a two-part tariff scheme that is feasible on the samples is also feasible on a new problem instance with high probability. We then show that computing revenue-maximizing feasible prices is hard even for buyers with additive valuations. Then, for bidders with identical valuation distributions, we present a condition that is sufficient for the two-part tariff scheme from the unsegmented setting to be optimal and feasible for the market-segmented setting. Finally, we prove a generalization result that states how many samples suffice so that we can compute the unsegmented solution on the samples and still be guaranteed that we get a near-optimal solution for the market-segmented setting with high probability.

1.3 Additional Related Research

(Menus of) two-part tariffs have been studied in economics [Feldstein, 1972; Ng and Weisser, 1974; Leland and Meyer, 1976; Murphy, 1977; Maskin and Riley, 1984; Wilson, 1993; Armstrong and Vickers, 2001; Sundararajan, 2004; Shi *et al.*, 2009]. The approach taken by much of the economic literature on this topic is rather different from the perspective we pursue: most work aims to find closed-form solutions for revenue maximizing two-part tariff menus, and in attempting to do so often places various (strong) restrictions on the setting. For example, Kolay and Shaffer [2003] derive closed forms for the profit-maximizing length-two menu of two-part tariffs when there are exactly two types of buyers. Bagh and Bhargava [2013] derive further closed-form results when valuations come from a finite discrete distribution. They moreover consider three-part tariffs—which has an additional quantity allowance after which the per-unit price takes effect. Schlereth et al. [2010] study some algorithmic aspects of finding revenue-maximizing TPTs. They cast the revenue-maximization problem as a mixed integer linear program and compare the performance of a few different heuristic solution algorithms. We can also write a mixed integer linear program to solve revenue maximization in our setting, but the algorithms we pose are more efficient. Other works consider two-part tariff pricing in relation to, for example, uncertainty [Lambrecht *et al.*, 2007; Png and Wang, 2010], opportunism [Marx and Shaffer, 2004],

and other practical buyer behavior [Narayanan *et al.*, 2007; Iyengar *et al.*, 2008]. To our knowledge, all prior work in economics considers continuous models, where quantity purchased is a continuous parameter and valuations are continuous and differentiable functions of quantity. Our setting considers a discrete and finite model, which is what gives rise to the interesting algorithmic challenges we tackle. In addition, the various examples of TPT pricing in the real world previously mentioned involve discrete quantities of goods, so our model is arguably a more realistic description of TPT pricing. TPTs have received some recent attention in computer science as well. Chawla and Miller [2016] study a form of TPTs (that is different from ours) in the context of finding simple mechanisms that yield (multiplicative) approximations to optimal revenue. However, they assume that the seller knows the distribution over buyers' values, and the mechanism design is tuned to that distribution. Notions of menu complexity and market segmentation have also been studied by computer scientists, though in different contexts [Babaioff *et al.*, 2017; Hart and Nisan, 2019; Cummings *et al.*, 2020].

To our knowledge, the only prior work that studied the model of TPTs that we address is that of Balcan et al. [2018]. However, they only studied sample complexity rather than algorithms. We take this a step further and solve the learning problem efficiently in terms of computation.

2 Algorithms for Optimal TPT Structures

In this section we study the computation of TPT structures that maximize empirical revenue over the given set of samples. We are given a set of samples $S = \{v^1, \ldots, v^N\}$, where each sample $v^i = (v_1^i(1), \ldots, v_1^i(K)), \ldots, (v_n^i(1), \ldots, v_n^i(K))$. That is, each sample gives a value for each buyer for each number of units bought. In the sample-based mechanism design literature, it is standard to assume a complete valuation draw like this in each sample. We also use the shorthand $v_j^i = (v_j^i(1), \ldots, v_j^i(K))$. In the first subsection we discuss computation of a single TPT and in the next subsection computation of a menu of multiple TPTs. In both sections we discuss the single-buyer case for simplicity, and then in the third subsection we present the generalization to the multi-buyer case.

2.1 An Efficient Algorithm for a Single TPT

In this subsection we give a polynomial-time algorithm to solve the empirical revenue maximization problem in the case where we can offer only one two-part tariff, that is, the menu length $L = 1$. Because in this section we are presenting the single-buyer case for simplicity, we do not include the buyer subscript in the valuations. So, our input is $S = \{v^1, \ldots, v^N\} = \{(v^1(1), \ldots, v^1(K)), \ldots, (v^N(1), \ldots, v^N(K))\}$.

We observe the following, which is key for our algorithm.

Lemma 1. *Suppose (p_1, p_2) is a TPT that maximizes empirical revenue over S. Then, the line with y-intercept p_1 and slope p_2 passes through a point $(q, v^i(q))$ for some $i \in \{1, \ldots, N\}$ and $q \in \{1, \ldots, K\}$.*

Proof. Consider a TPT line with y-intercept p_1 and slope p_2 that does not pass through any such point. Let $d > 0$

be minimal such that the TPT line with y-intercept $p_1 + d$ and slope p_2 passes through such a point. For any buyer j, $\mathrm{argmax}_q v^j(q) - (p_1 + p_2 q) = \mathrm{argmax}_q v^j(q) - (p_1 + d + p_2 q)$, so any buyer who was purchasing a nonzero quantity q when faced with (p_1, p_2) continues to purchase quantity q when faced with $(p_1 + d, p_2)$. Any buyer who purchased nothing when faced with (p_1, p_2) clearly continues to purchase nothing. Hence the TPT $(p_1 + d, p_2)$ generates strictly more revenue than (p_1, p_2). $\square$

Now, for each point $(q, v^i(q))$, we demonstrate that to find the revenue-maximizing tariff line it suffices to search over a set of at most NK tariff lines passing through $(q, v^i(q))$. Suppose we have a tariff line passing through $(q_0, v^{i_0}(q_0))$, at which buyer i buys quantity q_i. Let us compute the slope p_2 of such a tariff line at which i prefers to buy quantity q over q_i. We have

$$v^i(q) - (v^{i_0}(q_0) + p_2(q - q_0))$$
$$\geq v^i(q_i) - (v^{i_0}(q_0) + p_2(q_i - q_0))$$
$$\iff p_2 \geq (v^i(q_i) - v^i(q))/(q_i - q).$$

For i, q, q' let $\alpha_{i_0, q_0}(i, q, q') = \frac{v^i(q) - v^i(q')}{q - q'}$ denote the slope where buyer i switches preferences between quantity q and q' (this has no dependence on i_0 and q_0, but for notational consistency we leave the subscript). We must also account for the case where buyer i switches preferences between quantity q and purchasing nothing, for which a similar calculation gives $\alpha_{i_0, q_0}(i, q, 0) = \frac{v^i(q) - v^{i_0}(q_0)}{q - q_0}$.

The algorithm works by "hinging" a TPT at every point of the form $(q, v^i(q))$, and then increases the slope of the tariff line at increments determined by the α values—these increments represent thresholds at which some buyer starts buying a different quantity. As the slope of the hinged tariff line increases, the quantity purchased by a given buyer changes at most K times, and hence there are at most NK slope thresholds to check—between any two thresholds the average revenue is linear in the TPT since the quantities purchased are constant, and so the maximum revenue must be attained at one of these thresholds. See Figure 1 for an illustration. We write down the precise algorithm as Algorithm 1.

Theorem 1. *Algorithm 1 finds the single tariff (p_1, p_2) that maximizes empirical revenue over a sample set of size N in $O(N^3 K^3)$ time.*

Proof. That the algorithm finds the maximum revenue TPT follows from the fact that if $p_2 < p_2'$ are two consecutive slopes checked by the algorithm, the average revenue is linear as the slope varies between p_2 and p_2' (revenue is linear since, by construction, the quantities purchased by each buyer are constant for slopes varying between p_2 and p_2') and hence it suffices to compute revenue at the endpoints. Lemma 1 shows that the empirical revenue maximizer passes through a point $(q, v^i(q))$, and since Algorithm 1 checks all slope thresholds for TPTs passing through each such point, we are guaranteed to find the TPT yielding the maximum empirical revenue.

We now count the number of steps taken by Algorithm 1. Line 2 involves NK iterations, Line 4 involves N iterations,
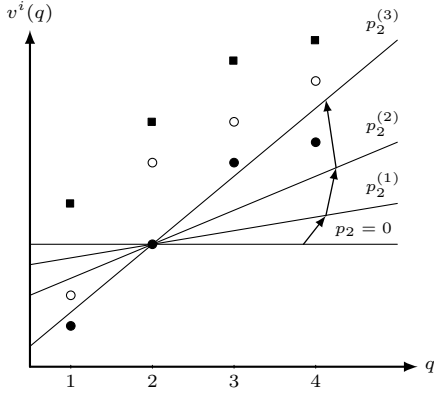
Figure 1: Three iterations of the single tariff algorithm from a given hinge point. The points displayed represent the valuations of three buyers (differentiated by the rendering style of the points) over four units. If, for example, $p_2', p_2'' \in (p_2^{(1)}, p_2^{(2)})$, then the quantities purchased by each buyer remain the same for the tariffs with slopes p_2' and p_2'' hinged at the given point.

Line 9 involves at most NK iterations (since each buyer starts by purchasing K units, and can change quantities at most K times as the slope p_2 increases), computing the minimization in Lines 10 and 11 requires at most NK steps, and computing revenue in Line 14 requires N steps. So the total run time is $O(NK(N + NK(NK + N))) = O(N^3 K^3)$. $\square$

We ran Gurobi (the fastest general-purpose mixed integer program solver) to find the revenue-maximizing single TPT for a single buyer (after formulating this problem as an integer program), and Algorithm 1 beat it dramatically. For example, averaged over 10 runs on randomly generated instances with $K = 5$ units and $N = 600$ samples, our algorithm returned the revenue-maximizing TPT in under 23 minutes while Gurobi took over 3.5 hours.

2.2 An Algorithm for Multiple TPTs

In this subsection we give an algorithm for optimizing a menu of two-part tariffs. In most applications, for practical reasons, the length of the menu cannot be very long, so L is a small constant (typically 2 or 3). We present an algorithm that is exponential in L but still polynomial in N and K. It can be viewed as a generalization of the single tariff algorithm. The geometric structure of the problem is the same as in Balcan et al. [2018], but we exploit it to get algorithms while they use it to prove sample complexity bounds.

Theorem 2. *There is an algorithm that finds the empirical revenue maximizing length L menu of tariffs over a sample set of size N in $(NK)^{O(L)}$ time.*

Proof. For input valuations $v^1, \ldots, v^N$, let $H_i(q, q', r, r')$ denote the hyperplane

$$v^i(q) - (p_1^r + q \cdot p_2^r) = v^i(q') - (p_1^{r'} + q' \cdot p_2^{r'}),$$

where if q (or q') is 0, the LHS (or RHS) is replaced by 0. Consider the collection of hyperplanes $\mathcal{H}$ consisting of these (at most $N(KL)^2$) hyperplanes for each $q, q' \in \{0, \ldots, K\}$, $r, r' \in \{1, \ldots, L\}$, $i \in \{1, \ldots, N\}$. It is a basic combinatorial fact that $\mathcal{H}$ partitions $\mathbb{R}^{2L}$ into at most $|\mathcal{H}|^{2L} \leq$

Algorithm 1 Single TPT for a Single Buyer

Input: Set of samples $S = \{v^1, \ldots, v^N\}$
Output: TPT (p_1, p_2) maximizing empirical revenue over S
1: Initialize MaxRev $\leftarrow 0$.
2: **for** $(i, q) \in \{1, \ldots, N\} \times \{1, \ldots, K\}$ **do**
3:      $p_2 \leftarrow 0$
4:      **for** $j \in \{1, \ldots, N\}$ **do**
5:          **if** $v^j(K) \geq v^i(q)$ **then**
6:              $q_j \leftarrow K$ (This makes the typical assumption that each buyer's valuation is nondecreasing in quantity. If that is not the case, one can in this line simply loop over quantities for each buyer separately in order to find the utility-maximizing quantity for that buyer.)
7:          **else**
8:              $q_j \leftarrow 0$
9:      **repeat**
10:          $p_2 \leftarrow \min_{j, q'} \alpha_{i,q}(j, q_j, q')$
11:          $j^*, q^* \leftarrow \operatorname{argmin}_{j, q'} \alpha_{i,q}(j, q_j, q')$
12:          $q_{j^*} \leftarrow q^*$
13:          $p_1 \leftarrow v^i(q) - q \cdot p_2$
14:          MaxRev $\leftarrow \max\{\text{MaxRev}, \widehat{\text{Rev}}_S(p_1, p_2)\}$
15:      **until** no update is found
16: **return** revenue maximizing (p_1, p_2)

$N^{2L} K^{4L}$ regions (each region is a connected component of $\mathbb{R}^{2L} \setminus \cup_{i,q,q',r,r'} H_i(q, q', r, r')$ and is an intersection of at most $|\mathcal{H}|$ halfspaces). The average revenue over the set of samples is linear within each such region, since the quantity purchased by each buyer remains constant within each region, so the maximum revenue within a region C can be found by solving the following linear program: if buyer i purchases quantity $q_i(C) \in \{0, \ldots, K\}$ of tariff $r_i(C) \in \{1, \ldots, L\}$ within C, the maximum revenue in C is

$$\max_{\mathbf{p} \in C} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(q_i(C) \geq 1) \cdot (p_1^{r_i(C)} + q_i(C) \cdot p_2^{r_i(C)}).$$

Each linear program involves $2L$ real variables, and $|\mathcal{H}| \leq NK^2$ constraints. So, it can be solved in $poly(N, K, L)$ time.

Moreover, there is a simple algorithm with run time $poly(N^{2L} K^{4L})$ which outputs a representation of each region determined by $\mathcal{H}$ as a $0/1$ vector of length $|\mathcal{H}|$, where the kth entry determines on which side of the kth hyperplane of $\mathcal{H}$ the region lies on. The high-level idea is to sequentially add each hyperplane to the list of regions maintained so far (starting with the entire Euclidean space $\mathbb{R}^{2L}$), iterating over the current regions and checking whether the added hyperplane intersects each region—updating the list of regions if so. See [Xu, 2020] for a more detailed description of enumerating the regions formed by a collection of hyperplanes (in a totally different context). Our algorithm solves the aforementioned linear program for every such region and picks the solution that yields highest empirical revenue. $\square$

In the single TPT ($L = 1$) case, Algorithm 1 is more efficient than the algorithm presented in Theorem 2. This is

because the former is a direct combinatorial algorithm that does not require solving LPs.

2.3 Generalization to Multiple Buyers

While our algorithms in the two subsections above were presented in the single buyer setting, they directly extend to the multi-buyer setting as follows.

Algorithm 1 generalizes by feeding in not just N valuations but all nN valuations. The only change to the algorithm itself is that in Step 12 we check that the allocation is feasible, that is, not more than the total quantity K is bought among the buyers in any sample $i \in \{1, \ldots, N\}$; if more is bought, then the assignment on that line is skipped because that pricing solution is infeasible. (Another nuance is that the initialization in Lines 2–8 might not be feasible, but that is fine.)

Similarly, in our multi-TPT algorithm, whenever we are about to solve an LP corresponding to some region determined by the set of hyperplanes, we first check that the region is feasible in the sense that the total quantity bought by buyers in any one sample is at most K.

Remark. Suppose buyers have additive valuations, that is, $v^i(q_1 + q_2) = v^i(q_1) + v^i(q_2)$ for any quantities $q_1 + q_2 \leq K$. Then, the revenue maximization problems considered in this section become trivial. In particular, the run time dependence on K and L vanishes. This can be seen due to the fact that $\text{price}(q) := \min_r p_1^r + q \cdot p_2^r$ is a piece-wise linear increasing concave function. An additive buyer's valuation function is simply a line with positive slope passing through the origin, as $v^i(q) = v^i(1 + \cdots + 1) = q \cdot v^i(1)$. Hence, the difference $v^i(q) - \text{price}(q)$ is always maximized when $q = K$, that is, buyers are always only interested in the entire K-unit bundle. Thus, revenue is determined by a single price, that of the entire bundle, and the seller simply can try every possible price in $\{v^1(K), \ldots, v^N(K)\}$, due to Lemma 1.

3 Market Segmentation

We now consider a setting in which each buyer belongs to one of M markets $\mathcal{X}_1, \ldots, \mathcal{X}_M$—determined by attributes such as geographic location, income level, etc. The seller sets M length L-menus of TPTs $(\mathbf{p}_1, \ldots, \mathbf{p}_M)$, where buyers in market m are allowed to purchase according to $\mathbf{p}_m$.

The seller wants to offer a TPT menu for each market so that the overall solution across markets is feasible, that is, that the sum of the demands of the markets does not exceed K. We show that any solution that is feasible for each sample in a large enough sample set is with high probability a feasible solution for any future sample.

Proposition 1. *Let $N \geq N_{TPT}(\varepsilon, \delta)$. With probability at least $1 - \delta$ over the draw of $S \sim \mathcal{D}^N$, if $\mathbf{p}_1, \ldots, \mathbf{p}_M$ is feasible for S,*

$$\Pr_{v \sim \mathcal{D}}[(\mathbf{p}_1, \ldots, \mathbf{p}_M) \text{ is feasible for } v] \geq 1 - \varepsilon.$$

Proof. Consider the class of $0/1$ valued indicator functions $\{f_v(\mathbf{p}_1, \ldots, \mathbf{p}_M)\}$ indicating whether $(\mathbf{p}_1, \ldots, \mathbf{p}_M)$ is feasible for v. For a single sample i, consider the set of hyperplanes of the form

$$v_j^i(q) - (p_{1,m(i,j)}^r + q \cdot p_{2,m(i,j)}^r) = v_j^i(q') - (p_{1,m(i,j)}^{r'} + q' \cdot p_{2,m(i,j)}^{r'})$$

for each j, q, q', r, r', where $m(i, j)$ is the market to which buyer j in sample i belongs. These hyperplanes partition the tariff space $\mathbb{R}^{2LM}$ into at most $(n(KL)^2)^{2LM}$ regions such that the indicator is constant within each region. Thus, by Theorems 2.2 and C.1 of Balcan et al. [2018], for a sample set S of size at least $N_{TPT}(\varepsilon, \delta)$, it holds with probability at least $1 - \delta$ that for all $\mathbf{p}_1, \ldots, \mathbf{p}_M$,

$$|\widehat{f}_S(\mathbf{p}_1, \ldots, \mathbf{p}_M) - \mathrm{E}_{v \sim \mathcal{D}}[f_v(\mathbf{p}_1, \ldots, \mathbf{p}_M)]| \leq \varepsilon.$$

So, if $(\mathbf{p}_1^S, \ldots, \mathbf{p}_M^S)$ is any feasible solution for S, that is, $\widehat{f}_S(\mathbf{p}_1^S, \ldots, \mathbf{p}_M^S) = 1$, we have

$$\mathrm{E}[f_v(\mathbf{p}_1^S, \ldots, \mathbf{p}_M^S)] = \Pr[f_v(\mathbf{p}_1^S, \ldots, \mathbf{p}_M^S) = 1] \geq 1 - \varepsilon,$$

with probability at least $1 - \delta$ over the draw of S and v. $\quad\square$

It turns out that market segmentation introduces substantial computational hurdles to revenue maximization. Even when bidders are additive (which removes the parameters L and K from the problem as remarked in the previous section), the problem of setting a feasible price for each market in an empirical revenue maximizing way is NP hard. Since additive buyers either purchase the entire bundle of K units or nothing, each menu is reduced to a single price, so the seller's problem is to set prices $p_1, \ldots, p_M$ for each market. Any solution must be feasible for the set of samples, which means at most one buyer can purchase the full K units in each sample.

Theorem 3. *Consider a set of samples $S = \{v^1, \ldots, v^N\}$ where each buyer belongs to one of M markets. Even if all buyers have additive valuations, there is no algorithm that finds feasible prices $p_1, \ldots, p_M$ that maximize empirical revenue over S in time polynomial in M and N, unless $\mathsf{P} = \mathsf{NP}$.*

Proof. We reduce from Maximum Weight Independent Set. Given an instance $G = (V, E)$ of Maximum Weight Independent Set (without loss of generality assume G has no isolated vertices), label the vertices $V = \{v_1, \ldots, v_n\}$, and let $w_i = \text{weight}(v_i)$. Let $p_i = \frac{w_i |E|}{\deg(v_i)}$. We will have n markets $\mathcal{X}_1, \ldots, \mathcal{X}_n$, corresponding to the vertices of G.

For each $(v_i, v_j) \in E$, we introduce a sample consisting of a buyer in market $\mathcal{X}_i$ with value p_i and a buyer in market $\mathcal{X}_j$ with value p_j (ensuring that no feasible pricing solution can simultaneously offer p_i to market $\mathcal{X}_i$ and p_j to market $\mathcal{X}_j$). So we have a total of $|E|$ samples. Clearly, any feasible revenue maximizing solution involves offering market $\mathcal{X}_i$ either price p_i, or something higher than p_i (so that no buyer in market $\mathcal{X}_i$ across any of the samples makes a purchase).

Our construction yields a one-to-one correspondence between independent sets in G and feasible n-tuples of prices: an independent set $I \subseteq V$ with weight $W = \sum_{v_i \in I} w_i$ corresponds to a pricing solution where if $v_i \in I$, market $\mathcal{X}_i$ is offered p_i, and if $v_i \notin I$, market $\mathcal{X}_i$ is offered something higher than p_i. For a vertex v_i in the independent set, there are precisely $\deg(v_i)$ samples containing a buyer in market $\mathcal{X}_i$ who makes a purchase at price p_i, so the average revenue obtained by the pricing solution corresponding to I is $\frac{1}{|E|} \sum_{v_i \in I} \deg(v_i) \cdot p_i = W$, by the choice of p_i. This completes the (clearly polynomial time) reduction. $\quad\square$

Remark. This hardness is inherent to the limited supply setting. If the seller has unlimited supply, and K is instead the maximum quantity any buyer is willing to purchase, we can find the empirical revenue maximizing market-segmented solution in $M(nNK)^{O(L)}$ time simply by running the procedure described in the previous section restricted to each market in turn. This finds the empirical optimum over each market, and without capacity constraints, this is a feasible solution and thus the optimal market-segmented solution as well.

3.1 Buyers with Identically Distributed Valuations

To circumvent the hardness of feasible empirical revenue maximization over worst case instances, we now study a setting where each buyer's valuation vector is drawn from the same distribution. Each market is of a certain prescribed size, and buyers are indistinguishable across markets. An example of a natural real-world market segmentation that potentially satisfies this is segmentation based on geographic location. For example, there may be no reason to believe that the average buyer in San Francisco values a gym-membership plan any differently than the average buyer in Pittsburgh. We show that under certain conditions, it is optimal to treat every buyer equally—regardless of whether they come from a large market or a small market. This immediately yields a simple algorithm in which we solve the non-segmented version of the problem, and reuse the solution for the segmented version.

Suppose there are a total of n buyers across markets, and an α_m fraction of these buyers belong to market $\mathcal{X}_m$. For simplicity, we assume that the seller receives zero revenue on instances on which the chosen solution is infeasible. Suppose the optimal solution $(\mathbf{p}_1^*, \ldots, \mathbf{p}_M^*) = \operatorname{argmax}_{(\mathbf{p}_1, \ldots, \mathbf{p}_M)} \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}_1, \ldots, \mathbf{p}_M)]$ satisfies the property that in expectation, buyers from market $\mathcal{X}_m$ contribute an α_m fraction of the total revenue. That is, $\mathrm{E}_v[\mathrm{Rev}_{v|\mathcal{X}_m}(\mathbf{p}_m^*)] = \alpha_m \cdot \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}_1^*, \ldots, \mathbf{p}_M^*)]$ for each m. In this case, we can reuse the non-segmented solution. For a randomly drawn v, let $\mathcal{F}$ denote the event that $f_v(\mathbf{p}_1^*, \ldots, \mathbf{p}_M^*) = 1$. We have that

$$
\begin{aligned}
\mathrm{E}_v[\mathrm{Rev}_{v|\mathcal{X}_m}(\mathbf{p}_m^*)] &= \mathrm{E}_v\left[\mathrm{Rev}_{v|\mathcal{X}_m}(\mathbf{p}_m^*) \mid \mathcal{F}\right] \cdot \Pr_v[\mathcal{F}] \\
&= \mathrm{E}_v\Big[\sum_{v_m \in \mathcal{X}_m} \mathrm{Rev}_{v_m}(\mathbf{p}_m^*) \mid \mathcal{F}\Big] \cdot \Pr_v[\mathcal{F}] \\
&= \alpha_m n \, \mathrm{E}_v[\mathrm{Rev}_{v_m}(\mathbf{p}_m^*) \mid \mathcal{F}] \cdot \Pr_v[\mathcal{F}],
\end{aligned}
$$

so $\mathrm{E}_v[\mathrm{Rev}_{v_m}(\mathbf{p}_m^*) \mid \mathcal{F}] = \frac{1}{n} \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}_1^*, \ldots, \mathbf{p}_M^*) \mid \mathcal{F}]$ for each m. Thus, we can set $\mathbf{p}_1^* = \cdots = \mathbf{p}_m^*$, and hence we only need to search for an optimal solution in the non-segmented case that we then offer to every market. The empirical revenue maximizing menu of TPTs $\mathbf{p}$ in the non-segmented case can be computed in $(nNK)^{O(L)}$ time, as in the previous section. Finally, we provide the generalization guarantee for using the unsegmented solution in the market-segmented case.

Theorem 4. *Let $N \geq N_{TPT}(\varepsilon, \delta)$. In the above setting, with probability at least $1 - \delta$ over the draw of $S \sim \mathcal{D}^N$,*

$$
\left|\widehat{\mathrm{Rev}}_S(\mathbf{p}, \ldots, \mathbf{p}) - \max_{(\mathbf{p}_1, \ldots, \mathbf{p}_M)} \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}_1, \ldots, \mathbf{p}_M)]\right| \leq 2\varepsilon,
$$

where $\mathbf{p}$ is the empirical revenue maximizing solution when all the markets are combined.

Proof. Let $\mathbf{p}^*$ denote the expected revenue maximizer in the non-segmented case, so by the previous discussion, $(\mathbf{p}^*, \ldots, \mathbf{p}^*)$ is also the expected revenue maximizer in the segmented case. Theorem C.1 of Balcan et al [2018] yields $|\widehat{\mathrm{Rev}}_S(\mathbf{p}) - \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p})]|, |\widehat{\mathrm{Rev}}_S(\mathbf{p}^*) - \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}^*)]| \leq \varepsilon$ with probability at least $1 - \delta$ over the draw of S. As $\widehat{\mathrm{Rev}}_S(\mathbf{p}) \geq \widehat{\mathrm{Rev}}_S(\mathbf{p}^*)$ and $\mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p}^*)] \geq \mathrm{E}_v[\mathrm{Rev}_v(\mathbf{p})]$, applying the triangle inequality yields the result. $\square$

4 Conclusions and Future Research

TPTs are a frequently-used pricing scheme in many applications. In this paper we studied efficient algorithms for finding revenue-maximizing TPT pricing schemes in the setting where the buyers' valuation distributions are given via samples. These algorithms have clear direct uses. They also provide the missing piece in the recent generalization theory for TPTs [Balcan *et al.*, 2018], which also uses samples of valuations as the input, and whose generalization approach is based on revenue-maximizing TPT schemes, but which did not provide any algorithms for computing such schemes.

We presented a polynomial algorithm for the case of optimizing one TPT. We also presented an algorithm for optimizing length-L menus of TPTs with complexity exponential only in L and polynomial in the other problem parameters.

We then generalized the problem of designing TPT schemes to multiple markets. We first proved how many samples suffice to guarantee that a TPT scheme that is feasible on the samples is also feasible on a new problem instance with high probability. We then showed that computing revenue-maximizing feasible prices is hard even for buyers with additive valuations. Then, for bidders with identical valuation distributions, we presented a sufficient condition for the TPT scheme from the unsegmented setting to be optimal and feasible for the market-segmented setting. Finally, we proved a generalization result that states how many samples suffice so that we can compute the unsegmented solution on the samples and still be guaranteed that we get a near-optimal solution for the market-segmented setting with high probability.

Several interesting open questions arise from our work. Can the $(NK)^{O(L)}$ algorithm in Theorem 2 be improved? In particular, can the exponential dependence on L be improved, or is the problem NP-hard? Moreover, the algorithm given in that theorem is fairly general purpose, and is not restricted to menus of TPTs. Can our method be used to develop efficient algorithms for other mechanism design settings? In the realm of market segmentation, are there interesting structural connections between the distributions over buyers' valuations in each market and potential polynomial time algorithms?

Acknowledgements

References

[Armstrong and Vickers, 2001] Mark Armstrong and John Vickers. Competitive price discrimination. *RAND Journal of Economics*, pages 579–605, 2001.

[Babaioff *et al.*, 2017] Moshe Babaioff, Yannai A Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proc. 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 869–877, 2017.

[Bagh and Bhargava, 2013] Adib Bagh and Hemant K Bhargava. How to price discriminate when tariff size matters. *Marketing Science*, 32(1):111–126, 2013.

[Balcan *et al.*, 2018] Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. A general theory of sample complexity for multi-item profit maximization. In *Proc. 2018 ACM Conference on Economics and Computation*, pages 173–174. ACM, 2018. Extended version on arXiv.

[Chawla and Miller, 2016] Shuchi Chawla and J Benjamin Miller. Mechanism design for subadditive agents via an ex ante relaxation. In *Proc. 2016 ACM Conference on Economics and Computation*, pages 579–596. ACM, 2016.

[Conitzer and Sandholm, 2002] Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *Proc. Eighteenth Conference on Uncertainty in Artificial Intelligence*, pages 103–110, 2002.

[Cummings *et al.*, 2020] Rachel Cummings, Nikhil R Devanur, Zhiyi Huang, and Xiangning Wang. Algorithmic price discrimination. In *Proc. Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2432–2451. SIAM, 2020.

[Feldstein, 1972] Martin S Feldstein. Equity and efficiency in public sector pricing: the optimal two-part tariff. *The Quarterly Journal of Economics*, pages 176–187, 1972.

[Hart and Nisan, 2019] Sergiu Hart and Noam Nisan. Selling multiple correlated goods: Revenue maximization and menu-size complexity. *Journal of Economic Theory*, 183:991–1029, 2019.

[Iyengar *et al.*, 2008] Raghuram Iyengar, Kamel Jedidi, and Rajeev Kohli. A conjoint approach to multipart pricing. *Journal of Marketing Research*, 45(2):195–210, 2008.

[Kolay and Shaffer, 2003] Sreya Kolay and Greg Shaffer. Bundling and menus of two-part tariffs. *The Journal of Industrial Economics*, 51(3):383–403, 2003.

[Lambrecht *et al.*, 2007] Anja Lambrecht, Katja Seim, and Bernd Skiera. Does uncertainty matter? Consumer behavior under three-part tariffs. *Marketing Science*, 26(5):698–710, 2007.

[Leland and Meyer, 1976] Hayne E Leland and Robert A Meyer. Monopoly pricing structures with imperfect discrimination. *The Bell Journal of Economics*, pages 449–462, 1976.

[Lewis, 1941] W Arthur Lewis. The two-part tariff. *Economica*, 8(31):249–270, 1941.

[Likhodedov and Sandholm, 2004] Anton Likhodedov and Tuomas Sandholm. Methods for boosting revenue in combinatorial auctions. In *AAAI*, pages 232–237, 2004.

[Likhodedov and Sandholm, 2005] Anton Likhodedov and Tuomas Sandholm. Approximating revenue-maximizing combinatorial auctions. In *AAAI*, pages 267–274, 2005.

[Marx and Shaffer, 2004] Leslie M Marx and Greg Shaffer. Opportunism and menus of two-part tariffs. *International Journal of Industrial Organization*, 22(10):1399–1414, 2004.

[Maskin and Riley, 1984] Eric Maskin and John Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.

[Murphy, 1977] Michael M Murphy. Price discrimination, market separation, and the multi-part tariff. *Economic Inquiry*, 15(4):587, 1977.

[Narayanan *et al.*, 2007] Sridhar Narayanan, Pradeep K Chintagunta, and Eugenio J Miravete. The role of self selection, usage uncertainty and learning in the demand for local telephone service. *Quantitative Marketing and Economics*, 5(1):1–34, 2007.

[Ng and Weisser, 1974] Yew-Kwang Ng and Mendel Weisser. Optimal pricing with a budget constraint–the case of the two-part tariff. *The Review of Economic Studies*, 41(3):337–345, 1974.

[Oi, 1971] Walter Y Oi. A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly. *The Quarterly Journal of Economics*, 85(1):77–96, 1971.

[Png and Wang, 2010] Ivan PL Png and Hao Wang. Buyer uncertainty and two-part pricing: theory and applications. *Management Science*, 56(2):334–342, 2010.

[Sandholm and Likhodedov, 2015] Tuomas Sandholm and Anton Likhodedov. Automated design of revenue-maximizing combinatorial auctions. *Operations Research*, 63(5):1000–1025, 2015.

[Schlereth *et al.*, 2010] Christian Schlereth, Tanja Stepanchuk, and Bernd Skiera. Optimization and analysis of the profitability of tariff structures with two-part tariffs. *European Journal of Operational Research*, 206(3):691–701, 2010.

[Shi *et al.*, 2009] Benlin Shi, Yang Song, Changchen Liu, and Yunfeng Luo. Competition and evolution of linear and two-part tariff. *Computers & Mathematics with Applications*, 57(11-12):1895–1900, 2009.

[Sundararajan, 2004] Arun Sundararajan. Nonlinear pricing of information goods. *Management Science*, 50(12):1660–1673, 2004.

[Wilson, 1993] Robert B Wilson. *Nonlinear pricing*. Oxford University Press on Demand, 1993.

[Xu, 2020] Haifeng Xu. On the tractability of public persuasion with no externalities. In *Proc. Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2708–2727. SIAM, 2020.