# Approximate Nonlinear Pricing*

Kristóf Madarász and Andrea Prat
STICERD, London School of Economics.

January 25, 2010

**Abstract**

Except in the one-dimensional case, single-agent mechanism design with quasi-linear payoffs is a computationally demanding problem. This paper proposes an approximation scheme based on categorization of types into 'stereotypes' and 'magnanimous' pricing, whereby the seller shares a small portion of the profits with the buyer. We show that, for any positive $\varepsilon$, our scheme finds (in polynomial time) a mechanism that generates an expected profit within an $\varepsilon$-factor of the optimal profit.

## 1 Introduction

In his path-breaking analysis of organizational decision-making, Herbert Simon argues that organizations recognize the limits imposed by our cognitive ability and develop institutions to achieve good results in the presence of such limits:

> "Most human-decision making, whether individual or organizational, is concerned with the discovery and selection of satisfactory alternatives; only in exceptional cases is it concerned with the discovery and selection of optimal alternatives. To optimize requires processes several orders of magnitude more com-

*We thank Mark Armstrong, Jim Minifie, Hervé Moulin and seminar participants at Cambridge University for useful suggestions.

> plex than those required to satisfice" (March and Simon, 1958, p 162).

When applied to a specific organizational problem, Simon's view spurs economists to ask two related questions. Given the cognitive constraints it faces, can the organization find the exact solution to this problem or will it settle for a good approximation? If so, will the features of the near-optimal solution differ systematically from those of the optimal one?

This paper attempts to answer these two questions in a very specific setting: single-agent mechanism design with quasilinear payoff. While this model has found numerous applications in economic, such as taxation and regulation, in what follows we identify it with its most common application: nonlinear pricing – namely the problem faced by a profit-maximizing monopolist who sells a product to a buyer, or a continuum of buyers. The monopolist can offer a menu of product specifications (quality or quantity) at different prices.

In the one-dimensional case of nonlinear pricing, there is a well-known way to characterize the optimal solution in a simple and powerful way by noting that downward local incentive-compatibility constraints must be binding (Mussa and Rosen 1978). However, most practical instances of nonlinear prices involve a multi-dimensional product space (even basic goods have numerous quality attributes) and a multi-dimensional type space (consumers with different income, location, age, etc). With more than one dimension, algorithms have been developed for special cases (Wilson 1993, Armstrong 1996). We also have an elegant characterization of the optimal solution in the general case (Rochet and Choné (1998), but it is not associated to a method for finding solutions. The hope of finding a computationally efficient general algorithm is slim, given that nonlinear pricing has been proven to be an NP-complete problem (Conitzer and Sandholm 2003).

Note that the difficulty of nonlinear pricing is of a strategic nature. Its computational complexity is not due to an intrinsic difficulty of determining the efficient allocation, but it hinges on the presence of an agency problem.[1] The problem is hard because the principal does not know the agent's type (informational asymmetry assumption) and the principal is maximizing profit not joint surplus (conflict of interest assumption). As we shall see, if either of these assumptions is dropped, the problem can be solved in polynomial time.

Given that finding an exact solution to nonlinear pricing seems hard, we turn our attention to approximations. We look for a method for finding solutions to nonlinear pricing that satisfies the following conditions: (i) It is not too computationally demanding; (ii) It yields an expected profit that is only marginally lower than the optimal profit; (iii) It satisfies (i) and (ii) for a general class of problems. Condition (iii) means that our algorithm will not necessarily be the best or simplest algorithm for certain specific classes of problems. But, in the spirit of Simon's view of organizational decision making, it will yield satisficing outcomes for a large class of problems.

The algorithm we propose is based on three steps.

1. Partition the type space into subsets of neighboring types. For each subset select one particular type to represent the whole subset. We call this type the *stereotype* of that subset.

2. Compute the optimal nonlinear pricing scheme for the set of stereotypes. As the computational time grows exponentially in the number of types (but not in the number of products), using stereotypes achieves a dramatic reduction in computation time.

3. Take the menu obtained in the second step: a vector of product-price pairs. Keep the product component unchanged and instead modify the price component as follows: offer a discount on each product that is proportional to the profit (revenue minus production cost) that the principal would get if she sold that product at the original price. The discount rate, which is dictated by the algorithm, depends on the

[1] This is in marked contrast with mechanism design with multiple agents, where the allocation problem is often intrinsically hard. For instance, even assuming that bidders report their valuations truthfully, winner determination is still NP-hard (de Vries and Vohra 2003).

number of stereotypes. It goes to zero as the stereotype partition becomes finer. In view of the last step, we call the algorithm, *Profit-Participation Pricing.*

We prove that Profit-Participation Pricing yields valid approximate solutions for a large class of nonlinear pricing problems. We only require that the number of dimensions of the type space be finite and that the utility function of the agent is continuous in his type. Specifically, the main result of this paper is that Profit-Participation Pricing is a Polynomial-Time Approximation Scheme (PTAS). Given any nonnegative number $\varepsilon$, the scheme returns a solution that yields a profit which is at least $1-\varepsilon$ of the profit generated by the optimal solution; the computation time is polynomial in the problem input.[2]

We can now tackle our second research questions. Are the features of our approximate solution systematically different from those of the optimal one? First, the number of products offered is lower and it is based on a rougher categorization of customers. As noted above, this is not due to an intrinsic difficulty of enumerating products, but to the computational cost of taking into account the agent's response. Second, the principal appears to "leave money on the table," in the sense that, because of the discounting method used, most products are not associated to a binding incentive-compatibility constraint or participation constraint, as it would happen in the exact solution. This pricing slack is of course a valid organizational response to the cost of computation, but it could appear 'magnanimous'. Indeed an outside observer who thought that the stereotype set chosen by the principal was the true type space rather than just an approximation would recommend to increase prices until each product is associated with a binding constraint. Similarly, if the observer would pick any set of types, the probability that he or she could find room for Pareto improvements is typically bounded away from zero.

[2] A more demanding notion of approximation quality is fully polynomial-time approximation scheme (FPTAS). After the main result, we conjecture that there is no FPTAS for nonlinear pricing.

Our approximation result has a mathematically equivalent interpretation in terms of sampling cost.[3] Suppose that the principal is not concerned with computation cost, but she does not know the agent's preferences. However, she can sample the type space. For a fixed marketing fee, she can observe the payoff function of a particular type. By incurring this sampling cost repeatedly, she can sample as many types as she wants. The Profit-Participation Pricing algorithm, as stated above, supplies the principal with an approximate solution whose total sampling cost is polynomial in the input size. In this interpretation, the principal first performs a market analysis leading to the identification of a limited set of typical consumers. Then, she tailors her product range to the stereotype set and prices it 'magnanimously' in the sense of Profit-Participation Pricing.

The paper is structured as follows. Section 2 introduces the nonlinear pricing model, discusses the computational complexity of finding an exact solution, and presents the notion of stereotype. Section 3 develops Profit-Participation Pricing and establishes an approximation bound (Lemma 1). Section 4 shows the main result of the paper, namely that the algorithm based on Profit-Participation Pricing is a PTAS (Theorem 4). Section 5 concludes by showing that our model has an equivalent interpretation in terms of a principal that does not know the agent's payoff function and by discussing future lines of research.

## 1.1 Literature

To the best of our knowledge, this is the first paper to provide a polynomial-time approximation scheme for a general class of single-agent mechanism design problems. The use of stereotypes in mechanism design and the idea of profit-participation pricing are also – we believe – original contributions of this paper. This brief section discusses the relation between our work and the relevant literature in economics and computer science.

There is a small but increasing economic literature which explicitly includes notions of computational complexity, such as Gilboa and Zemel

[3] This approach is related to Bergemann and Schlag (2008), who study the optimal pricing policy of a monopolist who faces model uncertainty.

(1989), Aragonès et al. (2005), Apesteguía and Ballester (2009), and Sher (2009), but none of these papers deal with mechanism design.

There is, however, economic work on approximation in mechanism design, motivated informally by complexity considerations. Armstrong (1999) studies near-optimal nonlinear tariffs for a monopolist as the number of product goes to infinity, under the assumption that the agent's utility is additively separable across products. He shows that the optimal mechanism can be approximated by a simple menu of two-part tariffs, in each of which prices are proportional to marginal costs (if agent's preferences are uncorrelated across products, the mechanism is even simpler: a single cost-based two-part tariff).[4]

In the context of multi-agent mechanism design, Bulow and Klemperer (1996) prove that, under certain conditions, an auction with $n$ bidders yields an expected revenue to the seller that is at least as large as *any* mechanism with $n-1$ bidders, implying that, as the number of potential buyers increases, auctions are a valid approximation for negotiations.[5]

A different formal definition of complexity that has been applied to mechanism design is communication complexity (Segal (2001), Nisan and Segal (2006), Segal (2007), Fadel and Segal (2009)). Rather than the time it takes to find the desired mechanism, the designer is concerned with the amount of information that must be communicated. In particular, approximation schemes are discussed in Nisan and Segal (2006). Our two approaches are complementary: while we study single-person mechanism design, communication complexity is most interesting in multiple-agent cases. In nonlinear pricing, communication complexity has limited bite as the communication burden of the optimal mechanism is linear in the minimum between the size of the type and the size of the product space.

A growing field of computer science, algorithmic mechanism design, approach mechanism design from a computational complexity perspective (see Hartline and Karlin (2007) for a survey). The area focuses on prior-free

[4] See also Chu, Leslie, and Sorensen (2009) for a study of bundling in the multi-product case, combining analytical, numerical, and empirical results.

[5] See also Hartline and Roughgarden (2009).

mechanisms, where the designer – as is the case in online mechanisms that must work for a range of environments – has no information on the agents' prior distributions. Instead, we have in mind designers, such as most firms, that have some information about their agents and make use of it when deciding on what mechanism to use.

An exception to the prior-free principle is Chawla, Hartline, and Kleinberg (2007). They study approximation schemes for single-buyer multi-item unit-demand pricing problems. The valuation of the buyer is assumed to be independently (but not necessarily identically) distributed across items. Chawla et al. find a constant-approximation algorithm based on virtual valuations (with an approximation factor of 3). Our paper differs in the type of approximation scheme that we choose, in the fact that we consider a general pricing problem, and in our search for an $\varepsilon$-approximation.

Some papers in management science study heuristics in the context of nonlinear prices (Green and Krieger 1985; Dobson and Kalish 1993). Their approximation bounds are numerical.

# 2 Model

## 2.1 Definitions

Consider a single-agent non-linear pricing problem. The principal offers a menu $M$ of product-price pairs to the agent. The agent must choose exactly one option from this menu. The set of available products is given by a finite set $Y$. The price of each option is some real number $p \in \Re$. We assume that the menu offered by the principal always contains an outside option $y_0, p_0$. The normalized price of $y_0$ is assumed to be $p_0 = 0$.

The agent's preferences depend on his type. The agent's type $t$ is an element of some finite set $T$. Although the set $T$ is common knowledge, the agent's actual type is his private information. The principal has a prior over $T$ which is given by a probability density function $f(t) \in \Delta T$. The fact that the agent's type is his private information is key in our model because it is the presence of such asymmetric information that results in severe computational complexity.

We assume quasi-linear payoffs. The agent's utility is his type-dependent valuation of the object net the price he pays to the principal:

$$v(t, y, p) = u(t, y) - p \tag{1}$$

The principal's profit is the price he receives for the object net the cost of producing the object:

$$\pi(t, y, p) = p - c(y) \tag{2}$$

Since both the type space $T$ and the product space $Y$ is multi-dimensional the agent's preference can greatly vary in his type. To constrain this variation, we impose two assumptions on the structure on the problem. First, we assume that the agent's type lives in some finite dimensional Euclidean space. Second, we assume that for any given multi-dimensional product $y$ the agent's preferences are Lipschitz continuous in his type.

Without loss of generality, we normalize the lower bound on $\Pi$ to equal 0. In short, a non-linear pricing problem is a tuple $\{T, Y, u, \pi\}$. We refer to the class of non-linear pricing problems that satisfy the above assumptions by $\Gamma$.

A solution to the principal's problem is a type-dependent allocation profile $y(t), p(t)$ that satisfies the agent's incentive compatibility constraints. We model this by assuming that the principal chooses a menu $M$ and offers this to the agent. The agent chooses an option from this menu and thus the resulting allocation profile $y(t), p(t)$ satisfies incentive compatibility (IC) by design. If $M$ also contains the outside option, the profile also satisfies the participation constraints (PC).

An optimal solution to the principal's problem is an allocation profile that satisfies (IC) and (PC) and maximizes the principal's expected profit. Formally, an optimal solution $y^*(t), p^*(t)$ to our non-linear pricing problem is a menu $M^*$ that maximizes:

$$\begin{aligned}\Pi(T, M^*) &= \sum_t f(t)\left[p^*(t) - c(y^*(t))\right] \qquad (3)\\ \text{such that } u(t, y^*(t)) - p^*(t) &\geq u(t, y^*(t')) - p^*(t') \text{ for all } t, t' \in T \\ \text{and } u(t, y^*(t)) - p^*(t) &\geq u(t, y_0) \text{ for all } t \in T\end{aligned}$$

## 2.2 Complexity of Finding an Exact Solution

Conitzer and Sandholm (Theorem 4, 2003) have already shown that the problem of findings an exact solution to single-agent mechanism design with quasilinear utility is NP-complete.

The intuition for this result is as follows. Under the Revelation Principle, we solve the problem in two stages: (i) For each possible allocation of products to types, we see if it is implementable and, if it is, we compute the profit-maximizing price vector; (ii) Given the maximized profit values in (i), we choose the allocation with the highest profit. While each step (i) is a linear program, the number of allocations that we must consider in (i) is as high as $(\#Y)^{\#T}$. The number of steps we must perform can then grow exponentially in the size of the input.[6]

This complexity result depends on two joint assumptions: asymmetric information and conflict of interest. If either of these assumptions is missing, we can find an exact solution in polynomial time.

If there were no asymmetric information and the principal could condition contracts on the agent's type, she would simply offer agent $t$ the surplus-maximizing allocation

$$y^*(t) \in \arg\max_y u(t, y) - c(y)$$

at price

$$p^*(t) = u(t, y^*(t)) - u(t, y_0)$$

[6] If there are less products than types, it may be quicker to compute an indirect mechanism rather than invoke the Revelation Principle and compute the direct mechanism. To achieve the exponential bound, assume for instance that $\#Y = a\#T$, where $a > 1$, and increase both the number of types and products.

This is would involve just $(\#Y)(\#T)$ steps.

If there were no conflict of interest – namely, if we wanted to maximize the surplus $u(t,y) - c(y)$ – it would be even simpler. The principal would offer all products, each of them at the cost of production ($p(y) = c(y)$). The agent would select $y^*(t) \in \arg\max_y u(t,y) - c(y)$. Solving this problem would involve just $\#Y$ steps.

## 2.3 Approximation Parameters

The quality of our approximation will depend on the parameters of the problem, which we define as follows.

**Type Topology** $T$ is an Euclidean metric space with finite dimensions, $\dim(T) \leq m \in \mathbb{R}$. $D$ is the size of the diagonal of the minimal hypercube that contains $T$.

**Lipschitz Continuity** There exists a number $K$ such that, for any $y \in Y$ and $t, t' \in T$

$$\frac{u(t,y) - u(t',y)}{d(t,t')} \leq K$$

Finally, we also assume that the principal's payoff is bounded: the upperbound is given by

$$\Pi_{\max} = \max_{t,y} u(t,y) - c(y) \tag{4}$$

To identify what constitutes an equivalence class of problems we note that taking positive affine transformation of the payoff functions will leave the approximation unaffected here. Hence the set of problems characterized by some $k$ where $k = K\frac{D}{\Pi_{\max}}$ is an equivalent class. In this fashion we can normalize $K = 1$ and $\Pi_{\max} = 1$ without loss of generality. The key parameters that determine the efficiency of the approximation will then be $D$ and $m$.Given the normalization and a fixed type space $T$ these two parameters determine how much 'variation' there is the agent's preferences.

## 2.4 Stereotypes

To reduce the complexity of the problem, we develop an approximation of the type space. Informally, we will partition the space of types into subset of nearby types. Then, for each type subset, we will choose a particular type – the *stereotype* – as representative of that subset.

Consider first a partition of the type space $T$. Let $\mathcal{P}(\varepsilon)$ denote a partition of $T$ such that the maximum distance between any two types $t, t' \in T$ and belonging to the same partition cell is $\varepsilon D$. Given $\mathcal{P}(\varepsilon)$, let $\tilde{f}_{\mathcal{P}}(t)$ denote the probability weight of the cell to which type $t$ belongs to for all $t$. In other words, $\tilde{f}_{\mathcal{P}}(t)$ is the sum of the probabilities of types $t'$ that belong to the same subset as type $t$. Although $\mathcal{P}(\varepsilon)$ is always a function of $\varepsilon$, in what follows we drop the $\varepsilon$-index from the notation.

Fixing a partition $\mathcal{P}$ , we pick one arbitrary stereotype as the representative for each cell of this partition. We call the collection of these stereotypes the stereotype set $S$. Let $\mathcal{S}(\mathcal{P})$ denote the collection of all stereotype sets for partition $\mathcal{P}$. In what follows, we will be looking for a partition $\mathcal{P}$ and a corresponding stereotype set $S$ such that the cardinality of the stereotype set is minimal while the partition still satisfies the $\varepsilon$ maximal distance property. Let $Q(\varepsilon)$ stand for the smallest cardinality of such a stereotype set $S$.To find an upper-bound on $Q(\varepsilon)$, let us partition the type space into identical $m$-dimensional hypercubes with diagonal length $\varepsilon D$. Given such a partition, the maximal number of stereotypes we need is:

$$\bar{Q}(\varepsilon) = \left(\frac{1}{\varepsilon}\right)^m \tag{5}$$

Note that this upper bound is tight is types are uniformly distributed on the type space and the number of types goes to infinity.

As an example, suppose $m = 2$, and $D = 10$. This upper bound means that for $\varepsilon = \frac{1}{2}$, we need at most 4 hypercubes to get a maximum distance of 5.

## 2.5 Stereotype Profit

An object that plays an important role in our algorithm is the profit that the principal could expect if the stereotype set $S$ was true. More precisely, let's consider the expected payoff of the principal when he offers menu $M$ and the type space is assume to be $S$ with probability weight $\tilde{f}_{\mathcal{P}}(t)$ on stereotype $t \in S$ . To determine this hypothetical profit we need to consider how stereotypes chooses from the offered menu. Given the incentive compatibility constraints for stereotype set $S$, we know that stereotypes choose such that

$$u(t, y(t)) - p(t) \geq u(t, y') - p' \text{ for all } (y', p') \in M \text{ and all } t \in S \tag{6}$$

Given the optimal behavior of the stereotypes, the hypothetical profit can be computed as

$$\Pi(S, M) = \sum_{t \in S} \tilde{f}_{\mathcal{P}}(t)(p(t) - c(y(t))) \tag{7}$$

Note, that there might be multiple profiles that satisfy the incentive compatibility constraint. Thus the above profit might not be unique. Although generically this will not be true, when it is, we define $\Pi(S, M)$ to be the maximal profit from the allocation profiles that satisfy the IC constraints of the stereotypes in $S$.

## 2.6 Failure of a Naive Approximation

One route towards finding an approximate solution is to select a stereotype set $S$ and find the optimal menu as if $S$ instead of $T$ was true. The menu that is optimal when the agent's true type is sampled only form $S$, can then be offered to the true type space.

However, this approach will not provide us with a near-optimal solution however. The reason is that there is no guarantee that types that are close to a stereotype will choose in similarly to how a stereotype does. Unless the type space is essentially one-dimensional, the binding constraint for a particular type may be non-local. In that case, it is possible to find a type near a stereotype that chooses an allocation that is far away from the allocation chosen by the stereotype, for instance by not participating. That in turn induces a type-specific profit that can be very different for the type under consideration and the original stereotype. In fact, Rochet and Choné show that 'bunching' is a robust feature of multi-dimensional nonlinear pricing (1998).

Let $t$ be a type nearby a stereotype $\hat{t}$ (let us write, somewhat intuitively, $t \sim \hat{t}$), and let $y\left(\hat{t}\right)$ and $y(t)$ be their chosen products respectively. We know from the line of reasoning above that it might not be that $y\left(\hat{t}\right) \sim y(t)$. It is true that Lipschitz continuity guarantees that the payoff of $t$ and $\hat{t}$ cannot be too far – else the one with the worse deal would just choose the allocation of the other one:

$$u\left(y\left(t\right), t\right) - p\left(y\left(t\right)\right) \sim u\left(y\left(\hat{t}\right), \hat{t}\right) - p\left(y\left(\hat{t}\right)\right).$$

However, this clearly does not imply that $p\left(y\left(t\right)\right) \sim p\left(y\left(\hat{t}\right)\right)$ – again, think of the case where $t$ does not participate and hence pays zero. Hence, it can be that $\pi\left(t, y\left(t\right)\right)$ is very different from $\pi\left(\hat{t}, y\left(\hat{t}\right)\right)$.

One conceivable way of guaranteeing that close types choose close allocations is to make more assumptions on $u$ and $c$. But those assumptions – if they exist – are likely to be very restrictive because bunching is known to occur even in very simple cases, such as the uniform-quadratic setting (Rochet and Choné 1998). In this paper, we follow an alternative route.

# 3 Profit Participation Pricing

In this section, we define the basic step in our approximation method, which we call Profit Participation Pricing and we prove one intermediate result. Both the definition and the result are somewhat general, because they will be applied twice – in different ways – in the proof of the main theorem.

Profit Participation Pricing is a modification of the naive method discussed above. Intuitively, take any menu (a vector of product-price pairs) given to a stereotype set. Keep the product component constant, just like in the naive method, but change the price component as follows: discount each original price by an amount that is proportional to the payoff that the principal receives for that particular product. In the approximation, the discounting coefficient will tend to zero. In other words, this scheme offers a profit-participation element to the agent.

Formally, let's fix the type space $T$, the set of allocations $Y$, and the players' preferences $u$ and $\pi$. Choose a stereotype set $S \in \mathcal{S}(\mathcal{P})$ that satisfies the $\varepsilon$-distance property. Now let's offer a menu $M$ for this hypothetical stereotype set and calculate the type-dependent allocation profile which results. Given optimal behavior by the agent, we can express the menu as $M = \left(\left(y^1, p^1\right), ..., \left(y^k, p^k\right)\right)$ where the superscripts identify the stereotypes. Alternatively, we can write this menu as $M = \cup_{\hat{t} \in S}\{y(t), p(t)\}$.

We now transform this menu $M = \cup_{t \in S}\{y(t), p(t)\}$ by holding the set of products fixed, but discounting the prices at which they are offered. In particular, let's define the discounted price of the option chosen by $\hat{t} \in S$, $\left(y(\hat{t}), p(\hat{t})\right) \in M$, to be

$$\tilde{p}\left(\hat{t}\right) = (1 - \tau)\, p\left(\hat{t}\right) + \tau c\left(\hat{t}\right) \tag{8}$$

This discounted price is equivalent to the price at which the principal loses $\tau$-fraction of the profit realized if stereotype $\hat{t}$ was true. Formally, $\tilde{p}\left(\hat{t}\right) - c\left(\hat{t}\right) = (1 - \tau)\left(p\left(\hat{t}\right) - c\left(\hat{t}\right)\right)$. The discounted menu thus contains the same products as the original one, but offers them at lower prices.

In what follows we will be focusing on a particular discounted menu where the fraction of the profit loss $\tau$ is set to equal:

$$\tau = \sqrt{2D\varepsilon} \tag{9}$$

Let's denote this specific discounted menu obtained by this Profit Participation Pricing by $\tilde{M} = \left(\left(y', \tilde{p}'\right), ..., \left(y^k, \tilde{p}^k\right)\right)$. The next lemma shows that when $\tilde{M}$ is offered to any stereotype set that is finer than $S$, which obviously includes the true type space $T$, the profit-loss to the principal is limited. By offering price discounts that are proportional to the profit generated by the stereotypes, the principal can ensure that even if types deviate from the behavior of their representative stereotypes, the impact of such deviations on the principal's profit will not be too large.

**Lemma 1** *Take any stereotype set $S \in \mathcal{S}(\mathcal{P})$ with minimum distance $\varepsilon$, and any menu $M$. Let $\tilde{M}$ be the menu derived through the Profit Participation Pricing. Take any stereotype set $S' \in \mathcal{S}(\mathcal{P}')$ where $\mathcal{P}'$ is a partition that is at least as fine as $\mathcal{P}$. Then:*

$$\Pi\left(S', \tilde{M}\right) - \Pi(S, M) \geq -2\sqrt{2D\varepsilon} \tag{10}$$

**Proof.** Take any menu $M$ and compute the discounted menu $\tilde{M}$. Consider two types $\hat{t}$ and $t$ that they belong to the same cell of $\mathcal{P}$ . For these two types it is always true that $\hat{t} \in S$ and $t \in S'$. We have to distinguish between two cases.

1. When $\tilde{M}$ is offered, $t$ chooses the allocation $y\left(\hat{t}\right)$ meant for $\hat{t}$. Here, the only loss for the principal is due to the price discount determined by $\tau$:

$$\tilde{p}\left(\hat{t}\right) - c\left(y\left(\hat{t}\right)\right) = (1-\tau)\left(p\left(\hat{t}\right) - c\left(y\left(\hat{t}\right)\right)\right)$$

2. When $\tilde{M}$ is offered, $t$ chooses the allocation $y'$ different from $y\left(\hat{t}\right)$. By the Lipschitz condition and the $\varepsilon$ distance limit we know that

$$\begin{aligned} \left|u\left(\hat{t}, y\left(\hat{t}\right)\right) - u\left(t, y\left(\hat{t}\right)\right)\right| &\leq D\varepsilon \\ \left|u\left(\hat{t}, y'\right) - u\left(t, y'\right)\right| &\leq D\varepsilon \end{aligned}$$

When combining these inequalities they imply that utility differentials for $t$ and $\hat{t}$ cannot be too different:

$$u\left(t, y\left(\hat{t}\right)\right) - u\left(t, y'\right) \geq u\left(\hat{t}, y\left(\hat{t}\right)\right) - u\left(\hat{t}, y'\right) - 2D\varepsilon$$

As the next step of the proof, we consider a revealed preference argument that helps determine .. Proof: IC's From the incentive compatibility constraints we know: (i) that $t$ prefers $y'$ to $y\left(\hat{t}\right)$ and hence

$$\tilde{p}\left(y\left(\hat{t}\right)\right)-\tilde{p}\left(y'\right)\geq u\left(t,y\left(\hat{t}\right)\right)-u\left(t,y'\right)$$

and (ii) that $\hat{t}$ preferred $\hat{y}\left(\hat{t}\right)$ to $y'$ and hence:

$$u\left(\hat{t},y\left(\hat{t}\right)\right)-u\left(\hat{t},y'\right)\geq p\left(y\left(\hat{t}\right)\right)-p\left(y'\right)$$

Now let's combine these inequalities with the Lipschitz bound and the definition of the discounted price $\tilde{p}$. We then get that:

$$\tau\left(p\left(y'\right)-c\left(y'\right)-\left(p\left(y\left(\hat{t}\right)\right)-c\left(y\left(\hat{t}\right)\right)\right)\right)\geq -2D\varepsilon$$

This inequality guarantees that a non-served type, $t\notin S$, will never choose an allocation $y'$ that is much worse than $y\left(\hat{t}\right)$ for the principal.

In total there are two sources of profit loss: (i) the price discount, (ii) the deviation of types from the chosen option of their stereotypes. The former is bounded by:

$$\pi\left(t,y',\tilde{p}\left(y'\right)\right)-\pi\left(t,y',p\left(y'\right)\right)\leq -\tau\pi\left(t,y',p\left(y'\right)\right)\leq -\tau$$

and the latter is bounded by:

$$\pi\left(t,y',p\left(y'\right)\right)-\pi\left(t,y\left(\hat{t}\right),p\left(y\left(\hat{t}\right)\right)\right)\leq -\frac{2D\varepsilon}{\tau}$$

To minimize total profit loss we can choose the discount rate $\tau$ such that for any type $t\in S'$,

$$\begin{aligned}\pi\left(t,y',\tilde{p}\left(y'\right)\right)-\pi\left(t,y\left(\hat{t}\right),p\left(y\left(\hat{t}\right)\right)\right) &\geq -\tau-\frac{2D\varepsilon}{\tau}\\ &= -2\sqrt{2D\varepsilon}\end{aligned}$$

This concludes the proof. ■

## 4 Near-Optimal Design

In the previous section, we introduced profit-participation pricing. We showed that PP allows one to adapt a menu from a stereotype set to a larger type set with a limited profit loss. The key insight was to introduce price discounts for the products on the menu in proportion to the principal's expected profit on these items. PP pricing guaranteed that the item a particular type chose from the menu generated a profit which was similar to the profit generated by its stereotype. More precisely, PP pricing provided a tight upper-bound on the principal's profit-loss when comparing the value of a menu offered to a stereotype set to the discounted version of this menu offered to a type set finer than the stereotype set.

In this section, we turn to how the PP pricing can be helpful in finding an approximate solution of the principal's problem. Note first that in the above discussion we did not talk about optimality. In this manner, neither the set of products offered to the agent nor the prices of these products satisfied any form of optimality. Let's now turn to the problem of optimality. To address this issue we now introduce a solution method that combines finding the optimal menu for a stereotype set and then adapting this menu to the true type set through PP pricing. We call this method the profit participation algorithm.

**PP Algorithm**

1. Given a problem from $\Gamma$ and an $\varepsilon > 0$, select a partition $\mathcal{P}$ on $T$ that satisfies the minimum distance property for this $\varepsilon$.

2. Select a stereotype set $S \in \mathcal{S}(\mathcal{P})$ and compute the **optimal** menu $\hat{M}$ for the problem $S, Y, u, \pi$.

3. Use $\hat{M}$ to obtain a discounted menu $\tilde{M}$ through the PP pricing.

The PP algorithm takes the pricing problem described in Section 2 as its input. It produces an output that consists of a menu $\tilde{M}$. Importantly, this output is a function of $\varepsilon$ the minimal distance property of the partition. To be precise, the output thus should be expressed as $\tilde{M}(\varepsilon)$. Note that since there are multiple ways to partition the type space $T$ multiple different stereotype sets $S$ to select on this partition for a fixed $\varepsilon$, $\tilde{M}(\varepsilon)$ is not generically not unique. Rather $\tilde{M}(\varepsilon)$ should be understood as a set of menus. Since the analysis below will hold for all members of this set, we refer to $\tilde{M}(\varepsilon)$ as any element of this set.

Given our normalization where $K = 1$ and $\Pi_{\max} = 1$, we say that a solution $\tilde{M}(\varepsilon)$ is within a factor $\varepsilon$ of the optimal solution if when this menu is offered to the whole type space, the principal's profit loss is at most $\varepsilon$. Formally,

$$\Pi(T, M^*) - \Pi\left(T, \hat{M}(\varepsilon)\right) \leq \varepsilon.$$

The next definition introduces the classic notion of a polynomial-time approximation scheme (PTAS).

**Definition 1** *Given a normalized class of optimization problems and an $\varepsilon > 0$, an algorithm is a polynomial-time approximation scheme (PTAS) if:*

*(AS) It returns a solution that is within a factor $\varepsilon$ of being optimal.*

*(PT) For every $\varepsilon$, the running time of the algorithm is a polynomial function of the input size*

Recall that the input size of our single-agent mechanism design problem was proportional to $|T| * |Y|$. Given this fact and the above definition, we can now state the main result of this paper.

**Theorem 1** *For all problems in $\Gamma$ and any $\varepsilon > 0$, the PP Algorithm is a PTAS.*

**Proof.** Step 1. Define the optimal mechanism

$$M^* = \arg\max_M \Pi(T, M).$$

to be set of allocations that maximizes the principal's expected profit subject to the IC constraints and contains the outside option. Let's denote this optimal profit by $\Pi(T, M^*)$. Note that $M^*$ and hence generically $\Pi(T, M^*)$ are unknown objects and they remain unknown in our approach. In fact, all the sets and menus in the proof remain unknown, except the ones found through PP.This does not mean, however, that we will never be able to bound the profit distance between these unknown menus and a carefully constructed menu $M$.

Step 2. Among all possible stereotype sets $\mathcal{S}(\mathcal{P})$, pick $S_{\max} \in \mathcal{S}(\mathcal{P})$ such that it maximizes the principal's profit once $M^*$ is offered. Formally,

$$S_{\max} \in \arg\max_{S \in \mathcal{S}} \Pi(S, M^*)$$

The principal's profit when the agent's type is restricted to $S_{\max}$ must be better than the optimal profit:

$$\Pi(S_{\max}, M^*) \geq \Pi(T, M^*)$$

Step 3. Now let's apply our approximation lemma. The input type spaces are $S_{\max} \in \mathcal{S}(\mathcal{P})$ and any fixed stereotype set $S$ from $\mathcal{S}(\mathcal{P})$. The input menu is $M^*$. Let's denote the menu obtained by profit-participation scheme (PP) by $M'$. From Lemma 1 we know that

$$\Pi\left(S, M'\right) - \Pi(S_{\max}, M^*) \geq -2\sqrt{2D\varepsilon}$$

Step 4. Take any stereotype set $S \in \mathcal{S}$ and pick the menu $\hat{M}$ that is optimal for that stereotype:

$$\hat{M} \in \arg\max_{M} \Pi(S, M)$$

By definition given stereotype set $S$, this menu $\hat{M}$ is better for the principal than using menu $M'$ which we defined in Step 3. Hence

$$\Pi\left(S, \hat{M}\right) \geq \Pi\left(S, M'\right)$$

Step 5. Let's apply our approximation lemma again where the input type spaces are $S$, $\not{T}$ and the input menu is $\hat{M}$. The output discounted menu is $\tilde{M}$. From Lemma 1 it follows that the bound on the profit-loss is:

$$\Pi\left(T,\tilde{M}\right)-\Pi\left(S,\hat{M}\right)\geq -2\sqrt{2D\varepsilon}$$

Summing up the above five steps:

$$\begin{aligned}
\Pi\left(T,M^{*}\right) &= [\text{max profit}] && \text{(Step 1)}\\
\Pi\left(S_{\max},M^{*}\right) &\geq \Pi\left(T,M^{*}\right) && \text{(Step 2)}\\
\Pi\left(S,M'\right) &\geq \Pi\left(S_{\max},M^{*}\right)-2\sqrt{2D\varepsilon} && \text{(Step 3)}\\
\Pi\left(S,\hat{M}\right) &\geq \Pi\left(S,M'\right) && \text{(Step 4)}\\
\Pi\left(T,\tilde{M}\right) &\geq \Pi\left(S,\hat{M}\right)-2\sqrt{2D\varepsilon} && \text{(Step 5)}
\end{aligned}$$

and hence the proft-loss due to using $\tilde{M}$ instead of the optimal $M^{*}$ is bounded by:

$$\Pi\left(T,\tilde{M}\right)\geq\ \Pi\left(T,M^{*}\right)-4\sqrt{2D\varepsilon}$$

We can now prove that the profit participation scheme is an approximation scheme (AS). This is true because

$$\lim_{\varepsilon\to 0}4\sqrt{2D\varepsilon}=0$$

To prove that PP is polynomial in time (PT), fix an $\varepsilon>0$ and note that the cardinality of the minimal stereotype set $S$ here is

$$\#S=\bar{Q}\left(\varepsilon\right)=\left(\frac{1}{\varepsilon}\right)^{m}$$

Thus, the total computation time of PP is proportional to the number of steps needed to compute the optimal mechanism for the stereotype set $S$. The Revelation Principle guarantees that this number is bounded above by

$$\#Y^{\#S}$$

Hence, for any given $\varepsilon$, the dimension of the stereotype space $\#S$ is fixed, and the computation time of PP is polynomial in the input size $\#Y * \#T$. ■

The computation time for the PP algorithm is of the order of

$$\#Y^{\#S}$$

and thus it is polynomial in the number of possible products, $\#Y$, independent of the number of types $\#T$, and exponential in the size of the stereotype space $\#S$. This means that our algorithm is particularly successful in reducing the complexity of the type space. Once the seller is satisfied with, say, a 1% profit loss, her computation cost is independent of the complexity of the type space.

A more stringent notion of approximation quality, fully polynomial-time approximation scheme (FPTAS), requires the computation time to be polynomial not only in the input size but also in the quality of the approximation, namely in $\frac{1}{\varepsilon}$. It is easy to see that this requirement fails here. A designer who wants to move from a 1% approximation to a 0.5% approximation, a 0.25% approximation, and so on, will face an exponentially increasing computation time.

Indeed, we conjecture that nonlinear pricing, as defined here, does not have an FPTAS. A "strongly NP-complete" has no FPTAS (Garey and Johnson, 1974). Conitzer and Sandholm (2003, Theorem 4) use a reduction to INDEPENDENT SET in order to prove that single-agent quasilinear-payoff mechanism design is NP-complete. However, INDEPENDENT SET is known to be strongly NP-complete.

## 5 Discussion and Conclusion

The interest in finding satisficing procedures for economic problems goes back to at least to the seminal work of Simon (1956). This paper offered an efficient way to approximate the optimal solution of a general class of nonlinear pricing problems. In our model, finding an optimal solution is difficult not due to the complexity of describing or enlisting option, but due to strategic reasons. Optimal decision-making is hard because the principal needs to fully understand how her choices affect the choices of the agent. When the monopolist faces uncertainty about the agent's type, understanding strategic responses is complex. How costly it is to obtain the optimal solution, as expressed in computational time, is increasing in such uncertainty. Our Profit Participation pricing allows to reduce the strategic complexity of the problem at a limited cost. Based on the understanding of a small set of stereotypes this scheme offers a solution that can approximate the optimal for any positive constant fraction of loss.

SOMETHING HERE ON ORGANIZATIONAL DECISION-MAKING

As mentioned in the introduction, our analysis has an alternative, and almost immediate, interpretation in terms of sampling cost. Suppose that the principal knows the set of possible types, $T$, and the set of possible products, $Y$, but does not know the payoff function of the agent: $u : T \times Y \to \Re$ (but she knows is that $u$ satisfies the Lipschitz continuity). The principal can choose to sample as many types as she wants, but each sampling operation entails a fixed cost $\gamma$. Sampling is simultaneous, not sequential. The principal chooses a sampling set $S$ ex ante. By equating the sampling set $S$ with the stereotype set, we can apply the PP Algorithm, as defined above. Theorem 4 guarantees that the resulting pricing scheme is an $\varepsilon$-approximation of the optimal pricing scheme.

# References

[1] José Apesteguía and Miguel A. Ballester. The Computational Complexity of Rationalizing Behavior. Working Paper, Universitat Pompeu Fabra, 2009.

[2] Enriqueta Aragonès, Itzhak Gilboa, Andrew Postlewaite, and David Schmeidler. Fact-Free Learning. *American Economic Review* 95(5), 1355–1368, 2005.

[3] Mark Armstrong. Multiproduct Nonlinear Pricing. *Review of Economic Studies* 64(1): 51–75, January 1996.

[4] Mark Armstrong. Price Discrimination by a Many-Product Firm. *Review of Economic Studies* 66, 151–168, 1999.

[5] Dirk Bergemann and Karl Schlag. Robust Monopoly Pricing. Cowles Foundation Discussion Paper 1527RR, 2008.

[6] Jeremy I. Bulow and Paul D. Klemperer. Auctions vs. Negotiations. *American Economic Review* 86(1): 180–194, March 1996.

[7] Shuchi Chawla, Jason D. Hartline, and Robert D. Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce (EC 2007)*, pp. 243-251, 2007.

[8] Chenghuan Sean Chu, Phillip Leslie, and Alan Sorensen. Bundle-Size Pricing as an Approximation to Mixed Bundling. Working paper, Stanford University, 2009.

[9] Vincent Conitzer and Tuomas Sandholm. Automated Mechanism Design: Complexity Results Stemming from the Single-Agent Setting. In *Proceedings of the 5th International Conference on Electronic Commerce (ICEC-03)*, pp. 17-24, Pittsburgh, PA, USA, 2003.

[10] Sven de Vries and Rakesh V. Vohra. Combinatorial Auctions: A Survey. *INFORMS Journal on Computing* 15(3), 284–309, Summer 2003.

[11] Gregory Dobson and Shlomo Kalish. Heuristics for Pricing and Positioning a Product Line Using Conjoint and Cost Data. *Management Science* 39(2), 160–175, February 1993.

[12] Ronald Fadel and Ilya Segal. The Communication Cost of Selfishness. *Journal of Economic Theory* 144:. 1895-1920, 2009.

[13] M. R. Garey and D. S. Johnson. "Strong" NP-Completeness Results: Motivation, Examples, and Implications. *Journal of the ACM* 25(3), 499–508, July 1978.

[14] Itzhak Gilboa and Eitan Zemel. Nash and Correlated Equilibria: Some Complexity Results. *Games and Economic Behavior* 1, 80-93, 1989.

[15] Paul E. Green and Abba M. Krieger. Models and Heuristics for Product Line Selection. *Marketing Science* 4(1): 1–19, Winter 1985.

[16] Jason D. Hartline and Anna Karlin. Profit Maximization in Mechanism Design. In *Algorithmic Game Theory*, Editors: Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vizarani. October 2007.

[17] Jason D. Hartline and Tim Roughgarden. Simple versus Optimal Mechanisms. In *Proceeding of the ACM Conference on Electronic Commerce*, July 2009.

[18] James March and Herbert Simon. Organizations. New York: Wiley, 1958.

[19] Michael Mussa and Sherwin Rosen. Monopoly and Product Quality. *Journal of Economic Theory* 18: 301–317, August 1978.

[20] Noam Nisan and Ilya Segal. The Communication Requirements of Efficient Allocations and Supporting Prices. *Journal of Economic Theory* 129(1): 192-224, July 2006.

[21] Jean-Charles Rochet and Philippe Choné. Ironing, Sweeping, and Multidimensional Screening. *Econometrica* 66(4), 783–826, July 1998.

[22] Herbert A. Simon. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69(1), 99-118, February, 1955.

[23] Ilya Segal. Communication Complexity and Coordination by Authority. Working paper, Stanford University, 2001.

[24] Ilya Segal. The Communication Requirements of Social Choice Rules and Supporting Budget Sets. *Journal of Economic Theory* 136: 341-378, September 2007.

[25] Itai Sher. Persuasion and Limited Communication. Working paper: University of Minnesota, 2009.

[26] Robert B. Wilson. *Nonlinear Pricing*. Oxford University Press, 1992.