

---

# Reaping the Benefits of Bundling under High Production Costs

---

Will Ma  
Columbia University

David Simchi-Levi  
Massachusetts Institute of Technology

## Abstract

It is well-known that selling different goods in a single bundle can significantly increase revenue. However, bundling is no longer profitable if the goods have high production costs. To overcome this challenge, we introduce a new mechanism, Pure Bundling with Disposal for Cost (PBDC), where after buying the bundle, the customer is allowed to return any subset of goods for their costs.

We provide two types of guarantees on the profit of PBDC mechanisms relative to the optimum in the presence of production costs, under the assumption that customers have valuations which are additive over the items and drawn independently. We first provide a distribution-dependent guarantee which shows that PBDC earns at least  $1 - 6c^{2/3}$  of the optimal profit, where  $c$  denotes the coefficient of variation of the welfare random variable.  $c$  approaches 0 if there are a large number of items whose individual valuations have bounded coefficients of variation, and our constants improve upon those from the classical result of Bakos and Brynjolfsson (1999) without costs.

We then provide a distribution-free guarantee which shows that either PBDC or individual sales earns at least  $1/5.2$  times the optimal profit, generalizing and improving the constant of  $1/6$  from the celebrated result of Babaioff et al. (2014). Conversely, we also provide the best-known upper bound on the performance of any partitioning mechanism (which captures both individual sales and pure bundling), of  $1/1.19$  times the optimal profit, improving on the previously-known upper bound of  $1/1.08$ .

Finally, we conduct simulations under the same playing field as the extensive numerical study of Chu et al. (2011), which confirm that PBDC outperforms other simple pricing schemes overall.

## 1 Introduction

We study the monopolist pricing problem of a firm selling  $n$  different items to a single random customer from the population. For each item, the customer wants at most one copy, and has a *valuation*, or maximum willingness-to-pay, drawn from a known distribution. The firm offers take-it-or-leave-it prices for every subset of items, and the customer chooses the subset maximizing her *surplus*, assumed to equal to the *sum* of her valuations for the items in the subset minus the subset's price. Ties are broken in the firm's favor. The objective of the firm is to maximize expected revenue, with the restriction that the empty subset must be priced at 0.

In the full generality of the problem, the firm has  $2^n - 1$  prices to set. However, it is important to find profitable yet *simple* pricing schemes that are explained by a small number of prices. Two such schemes are *Pure Components* (PC), where items are priced separately and the price of a subset is understood to be the sum of its constituent prices, and *Pure Bundling* (PB), where the only option is to buy all of the items together at a fixed price. A third scheme that captures both PC and PB is *Mixed Bundling* (MB), which prices items individually, but offers a bundle discount if all of the items are purchased.

Simple pricing schemes such as PC, PB, and MB are ubiquitous in practice, and consequently their efficacy is a subject of great interest. When  $n = 1$ , the firm's only option is to sell the item individually, and the optimal individual price is the  $p$  which maximizes  $p(1 - F(p))$ , where  $F$  is the CDF of the item's valuation. However, when  $n > 1$ , bundling can often be better than individual pricing. For example, suppose there are two products with IID valuations, each of which is 1 w.p.  $1/2$ , and 2 w.p.  $1/2$ . If the items are sold

individually, then the firm can earn at most 1 from each item in expectation, for a total revenue of 2. On the other hand, if the items are sold in a bundle with price 3, then the expected revenue would be  $\frac{9}{4}$ .

The key observation is that the valuation of the bundle is more concentrated around its mean than the valuation of the individual items, allowing the firm to set a price that is the maximum willingness-to-pay for a larger fraction of customers. This power of PB over PC was first observed in the pioneering work of Stigler (1963), Adams and Yellen (1976), Schmalensee (1984), and McAfee et al. (1989). However, bundling is not always more profitable than individual sales, especially once production costs are considered. Taking the same example, suppose now that each item has an instantaneous production cost of 1.5. Selling the items individually at price 2 each yields a profit of  $(2 - 1.5) \cdot 1/2 = 1/4$  per item, for a total of  $1/2$ . On the other hand, the optimal bundle price is 4, which sells w.p.  $1/4$ , for a total profit of  $(4 - 3) \cdot 1/4 = 1/4$ . PB was no longer better here because the firm had to charge a high bundle price to cover its costs, and hence the customer needed to have high valuations for *both* items in order to make a purchase.

Over the decades, a lot of work has been done to compare the profit of PB vs. PC. Adams and Yellen (1976) write, “The chief defect of Pure Bundling is its difficulty in complying with Exclusion,” where Exclusion refers to the principle that a transfer is better off not occurring when the consumer’s valuation is below the producer’s cost. It is observed in Schmalensee (1984) for the case of bivariate normal valuations that PB is better when mean valuations are high compared to costs. Bakos and Brynjolfsson (1999) prove that bundling a large number of goods can extract almost 100% of the total welfare, but this is crucially dependent on the items being “information goods”, i.e. goods with no production costs. Fang and Norman (2006) characterize conditions under which PB outperforms PC for a fixed number of items, again highlighting the importance of low costs.

The indisputable conclusion from all this work is that high costs are the greatest impediment to the power of bundling. However, in this paper we argue that firms can reap the benefits of bundling even under high production costs. We propose a new pricing scheme called *Pure Bundling with Disposal for Cost* (PBDC), where all of the items are sold as a bundle, but the customer is then allowed to *return any subset of items* for a refund equal to their total production cost.

PBDC results in strictly higher consumer surplus than PB (with the same bundle price), because the customer can return items valued below cost for a refund

equal to cost. Meanwhile, the firm is indifferent between producing an item for its cost or returning its cost to the customer, but in the end the return option granted by PBDC entices more customers to buy the bundle, so the firm’s profit is strictly increased as well.

Logistically, the extra step in PBDC of returning products and processing the refund results in additional overhead costs, which are not captured in our model. However, the firm could enforce that the set of products to “return” must be decided at time of checkout, and consequently not purchased in the first place, to avoid these overhead costs. We believe that our presentation of PBDC as a pure bundle with a return option both helps consumers choose which subset to purchase, and also helps firms analyze the profit of bundling under costs. We now elaborate on the latter.

### 1.1 Outline of Theoretical Results

We analyze the profit of pricing according to PBDC (with an optimized bundle price) in comparison to the optimal pricing<sup>1</sup> facing production costs. If all these costs are zero, then the return option is inconsequential and PBDC coincides with the classical Pure Bundling. In this case, our results generalize the existing results for PB, and moreover lead to *improved guarantees* relative to the optimal pricing.

We also emphasize that generalizing from PB to our setting of PBDC with costs does not trivially follow from analyzing the items’ valuations shifted by their costs. This is because these shifted valuations could be negative, whereas existing analyses assume that valuations are non-negative. Truncating negative valuations to zero could unfortunately increase the optimum against which we are comparing, since the optimal revenue is generally *non-monotone* in the valuations (Hart and Reny, 2015). Consequently, an important part of our analysis is to account for the increase in the optimum from having negative valuations.

Our first result says that the profit of PBDC is at least  $1 - 6\nu_w^{2/3}$  times that of the optimum, where  $\nu_w$  denotes the coefficient of variation of the welfare random variable. This guarantee builds upon the results of Bakos and Brynjolfsson (1999); Armstrong (1999) which say that PB extracts nearly 100% of the optimum if there are a large number of items with independent valuations. Indeed, in such a regime  $\nu_w$  tends to 0 and our result says that PBDC is asymptotically optimal. Our analysis also leads to improved constants in the convergence rate (see Section 2 for details).

<sup>1</sup>Technically we are comparing to the optimal randomized mechanism, which is also allowed to set prices for lotteries over items (details in Section 3).

Our second result says that the profit of either PBDC or PC (with optimized prices) is at least 1/5.2 times that of the optimum. This guarantee builds upon the line of work by Hart and Nisan (2017); Li and Yao (2013); Babaioff et al. (2020) culminating in the statement that the revenue of either PB or PC is at least 1/6 times that of the optimum when there are no costs (they also show that the inclusion of PC in this statement is necessary). We improve the bound of Babaioff et al. (2020) from 1/6 to 1/5.2 by showing that worst cases for the *core* and the *tail* in their decomposition cannot occur simultaneously (see Section 3 for details).

Our final result is an example on which neither PBDC nor PC can earn more than  $\frac{3+\ln 2}{3+2\ln 2} \approx \frac{1}{1.19}$  times the optimum. In fact our example contains two IID items without costs, so it applies to the classical PB setting and improves the previous-best upper bound of  $\frac{12}{13} \approx \frac{1}{1.08}$  from Hart and Nisan (2017). We should note that a upper bound of  $\frac{1}{2}$  can be found in Rubinstein (2016), but in his example PB is actually optimal if one is allowed to *partition* the items into bundles, whereas our example provides an upper bound even on the partitioning strategy (see Section 3.1 for details).

## 1.2 Summary of Numerical Experiments

We repeat the numerical experiments from Chu et al. (2008), on the same valuation distributions and costs, with PBDC added in as a pricing scheme to be compared to PC, PB, and the *Bundle-Size Pricing* (BSP) they introduce<sup>2</sup> which achieves over 99% of the optimum in their simulations. BSP is defined by parameters  $P_k^{\text{BSP}}$  which indicate the price for taking *any* subset of size  $k$ , for all  $k = 1, \dots, n$ . Note that when items have identical costs (something common in the experimental settings), PBDC is a subfamily of BSP.

Nonetheless, our experiments show that PBDC still attains between 97.5% and 100% of the (nearly optimal) BSP profit in these settings. On the other hand, if costs are allowed to vary at all, then PBDC becomes the best-performing pricing scheme by far. In fact, the worst case for PBDC is the aforementioned setting where it attains 97.5%; contrast this with 79.9%, 16.8%, 59.5% for PC, PB, BSP respectively in their worst-case settings. In addition to being profit-maximizing, PBDC also achieves excellent total surplus in our simulations, and scales well with the number of items (see Section 4 for details).

All in all, since BSP contains PBDC as a subfamily

<sup>2</sup>Prior to Chu et al. (2008), the BSP pricing scheme has appeared as “Customized Bundling” in Hitt and Chen (2005). We closely follow the experimental parameters from the working paper Chu et al. (2008) but should note that the published version is Chu et al. (2011).

in the special case of identical costs, our theoretical guarantees for PBDC provide the first technical explanation for some of the empirical successes for BSP found by Chu et al. (2008). On the other hand, our experiments show that it is possible to outperform BSP on the instances in Chu et al. (2008) where costs are heterogeneous. Finally, we remark that this is all despite PBDC being faster to optimize over (having 1 parameter instead of  $n$ ) and simpler to interpret. One could also consider an extension of PBDC where each item’s return refund is optimized, instead of pegged to its production cost. However, such an optimization problem is generally non-trivial (cf. Li et al., 2020). We leave its solution as future work that would be a further improvement of PBDC.

## 1.3 Further Related Work

Simple mechanisms and bounds on their performance, in the special case of a single buyer, has been an active area of research over the past decade (Hart and Nisan, 2017, 2013; Hart and Reny, 2015; Li and Yao, 2013). Our distribution-free lower bound of 1/5.2 improves the celebrated 1/6-guarantee originally proved by Babaioff et al. in 2014. There has since been many generalizations of their result, to multiple buyers (Yao, 2015), sub-additive buyers (Rubinstein and Weinberg, 2018), buyers with common-base-value (Bateni et al., 2015) or proportional (Cai et al., 2019) complementary valuations, among others, which are described in Babaioff et al. (2020). These papers consider setting which are more general settings than ours in some ways, but to our knowledge, none of them consider costs, or have a guarantee better than ours of 1/5.2.

Outside of mechanism design, bundling is a very broad topic whose study was pioneered by Stigler (1963); Adams and Yellen (1976); Schmalensee (1984); McAfee et al. (1989). The problem of computing optimal bundle prices is addressed in Wilson (1993). Bundling can also have an impact on supply chain fulfillment, the way in which goods are marketed, for which we respectively refer to Ernst and Kouvelis (1999); Venkatesh and Mahajan (2009) and the reference therein.

## 1.4 Preliminaries

A firm has  $n$  different items for sale. For each  $i$ , the cost incurred by the firm for selling item  $i$  is  $c_i \geq 0$ .  $c_i$  can be thought of as an instantaneous production cost, the opportunity cost of saving the inventory for someone else, or the value of the item to the seller.

Each of the firm’s customers has a valuation vector  $x \in \mathbb{R}_{\geq 0}^n$  for the items. A customer wants at most one of each item, and her utility for a subset of items  $S$  is  $\sum_{i \in S} x_i$ .  $x$  is a random vector drawn from a distri-

bution  $D$  representing the valuation vectors across the population. The valuations are said to be *independent* if  $D$  is a product distribution  $D_1 \times \dots \times D_n$ , with  $x_i$  drawn independently from  $D_i$  for all  $i = 1, \dots, n$ . The firm's objective is to maximize its expected profit.

The firm sells the items by posting for every  $S \subseteq \{1, \dots, n\}$  the price  $P(S) \geq 0$  that must be paid to receive exactly the subset of items in  $S$ , with  $P(\emptyset) = 0$ . A customer with valuation vector  $x$  purchases a surplus-maximizing subset  $S^* \in \operatorname{argmax}_S (\sum_{i \in S} x_i - P(S))$ , in which case the firm earns profit  $P(S^*) - \sum_{i \in S^*} c_i$ . Ties are broken in the firm's favor.

A *pricing scheme* is a restriction on the complexity of the price function  $P$ , by forcing it to be defined by a small number of prices in a way that is simple for the customer to understand. We now recap some pricing schemes from the literature:

1. Pure Components (PC): the items have individual prices  $P_1^{\text{PC}}, \dots, P_n^{\text{PC}} \geq 0$ , and  $P(S) = \sum_{i \in S} P_i^{\text{PC}}$  for any subset  $S$ .
2. Pure Bundling (PB): there is a single bundle price  $P^{\text{PB}} \geq 0$ , and  $P(S) = P^{\text{PB}}$  for any  $S \neq \emptyset$ , thereby making the grand bundle the only viable purchase option.
3. Bundle-Size Pricing (BSP): there are prices  $P_1^{\text{BSP}} \leq \dots \leq P_n^{\text{BSP}}$  based on the number of items purchased, and  $P(S) = P_{|S|}^{\text{BSP}}$  for any  $S \neq \emptyset$ .

We introduce the following pricing scheme in this paper, which takes costs into account:

4. Pure Bundling with Disposal for Cost (PBDC): there is a price  $P^{\text{PBDC}} \geq 0$  for the grand bundle, and  $P(S) = P^{\text{PBDC}} - \sum_{i \notin S} c_i$  for any  $S \neq \emptyset$ , thereby returning the costs of items  $i$  not in  $S$  to the customer.

A price function  $P$  which can be described in the form of one of these pricing schemes is said to *fall under that pricing scheme*. The *profit of a pricing scheme* is then defined as the maximum expected profit that can be earned under the restriction that  $P$  must fall under that pricing scheme. In this paper we provide guarantees on the profit of PBDC, relative to the optimal unrestricted  $P$ , which hold over all instances. We now define these benchmarks on a given instance.

**Definition 1.1.** A *problem instance* is defined by costs  $c_1, \dots, c_n$  and a distribution  $D$  from which valuation vector  $x$  is drawn. Given an instance, define:

1. The *welfare* random variable to be  $w = \sum_{i=1}^n \max\{x_i - c_i, 0\}$ ;

2.  $\mu_w$  and  $\sigma_w$  to respectively denote the mean and standard deviation of  $w$ ;
3.  $\nu_w = \sigma_w/\mu_w$  to denote the coefficient of variation of  $w$ , assuming  $\mu_w > 0$  and  $\sigma_w < \infty$ ;
4.  $\text{OPT}$  to denote the maximum expected profit that could be earned by any price function  $P$  satisfying  $P(S) \geq 0 \forall S$  and  $P(\emptyset) = 0$ .

In Section 2 we provide guarantees on the profit of PBDC relative to  $\mu_w$ , which *depend on  $\nu_w$* . Note that  $\mu_w$  is easily shown to be an upper bound on  $\text{OPT}$  (see e.g. Bakos and Brynjolfsson, 1999). In Section 3 we provide *distribution-free* guarantees on the profit of either PBDC or PC relative to the optimal randomized selling mechanism (which is a tighter upper bound on  $\text{OPT}$ ). Note that in the special case where all costs  $c_i = 0$ , PBDC is equivalent to PB, and welfare  $w = \sum_{i=1}^n x_i$ . As a result, our guarantees for PBDC generalize existing results for PB and existing results on the profit of either PB or PC.

## 2 Distribution-dependent Guarantees

Our main result in this section is Theorem 2.1 and its Corollary 2.2. Our analysis uses Cantelli's inequality and the weighted arithmetic mean-geometric mean inequality, which are described in references Lugosi (2009) and Zhao (2008) respectively. All proofs are deferred to the Arxiv version of this paper (Ma and Simchi-Levi, 2015).

**Theorem 2.1.** For all  $\varepsilon \in [0, 1]$ , the expected revenue of PBDC with bundle price  $P^{\text{PBDC}} = (1 - \varepsilon)\mu_w$  is at least  $\frac{\varepsilon^2 - \varepsilon^3}{\varepsilon^2 + \nu_w^2} \cdot \mu_w$ . In particular, if

$$\varepsilon = \frac{2\nu_w^{2/3}}{3\nu_w^{2/3} + 2}, \quad (1)$$

then the expected revenue of PBDC is at least

$$\frac{4\mu_w}{4 + 24\nu_w^{2/3} + 45\nu_w^{4/3} + 27\nu_w^2} \quad (2)$$

which in turn is at least

$$(1 - 6\nu_w^{2/3}) \cdot \mu_w. \quad (3)$$

Lower bound (3) shows that for small  $\nu_w$ , the fraction of  $\mu_w$  which PBDC is able to extract is  $1 - \Theta(\nu_w^{2/3})$ , recovering an existing result from Bakos and Brynjolfsson (1999) and Armstrong (1999). However, for large  $\nu_w$ , our tighter bound (2) still provides a non-trivial revenue guarantee, something not accomplished in the existing analyses. Interestingly, from (1) we can see

that the bundle price  $P^{\text{PBD}}$  should not be set lower than  $\mu_w/3$ . This is a useful managerial reference point in situations where  $\mu_w$  is known but the exact demand distribution  $D$  is not (Chen et al., 2019), in which case the optimal value for  $P^{\text{PBD}}$  cannot be computed.

Theorem 2.1 provided a guarantee based on the coefficient of variation of the welfare random variable  $w$ , with the guarantee approaching 100% as  $\nu_w \rightarrow 0$ . We now see that if there are a large number of items  $n$  with *independent* valuations, then indeed the law of large numbers ensures  $\nu_w$  to be small.

**Corollary 2.2.** *Suppose that  $x_1, \dots, x_n$  are independent. Let  $\mu_{\min}$  be a uniform lower bound on the mean of  $\max\{x_i - c_i, 0\}$  of all  $i$ , and let  $\sigma_{\max}^2$  be a uniform upper bound on the variance of  $\max\{x_i - c_i, 0\}$  of all  $i$ , with  $\mu_{\min} > 0$ ,  $\sigma_{\max} < \infty$ , and  $n > (\frac{\sigma_{\max}}{\mu_{\min}})^2$ . Then the revenue of PBDC is at least*

$$(1 - 6(\frac{\sigma_{\max}}{\mu_{\min}})^{2/3} \frac{1}{\sqrt[3]{n}}) \cdot \mu_w.$$

Taking  $n \rightarrow \infty$  with  $\mu_{\min}$  and  $\sigma_{\max}$  fixed, we see that PBDC extracts the entire welfare. Note that truncating  $x_i - c_i$  from below by 0 can only increase the mean and decrease the variance, so any bounds on the mean and variance of the untruncated  $x_i - c_i$  also suffice for Corollary 2.2.

### 3 Distribution-free Guarantee

Our main results in this section are Theorems 3.2 and 3.4 analyzing the profit of PBDC relative to the optimal randomized mechanism, which we now define. All proofs are deferred to the Arxiv version of this paper (Ma and Simchi-Levi, 2015).

**Definition 3.1.** For any costs  $c_1, \dots, c_n$  and valuation distribution  $D$  supported on  $\mathcal{X}$ , define  $\text{REV}(D)$  to be the optimal objective value of the following problem:

$$\begin{aligned} & \max \mathbb{E}_{x \sim D} \left[ s(x) - \sum_{i=1}^n c_i q_i(x) \right] \\ \text{s.t. } & \sum_{i=1}^n x_i q_i(x) - s(x) \geq \sum_{i=1}^n x_i q_i(y) - s(y) \quad \forall x, y \in \mathcal{X} \\ & \sum_{i=1}^n x_i q_i(x) - s(x) \geq 0 \quad \forall x \in \mathcal{X} \\ & q(x) \in [0, 1]^n \quad \forall x \in \mathcal{X} \end{aligned}$$

In the LP,  $q_i(x)$  denotes the probability of a customer with valuation vector  $x$  receiving item  $i$ , for all  $x$  and  $i$ , while  $s(x)$  denotes the expected total price paid. The first constraints impose that a customer with valuation vector  $x$  cannot receive greater surplus from lying

about her valuation vector being  $y$  (this is known as *incentive-compatibility*). The second constraints impose that all customers receive non-negative surplus (this is known as *individual-rationality*).

Any feasible mechanism is equivalent to offering a menu consisting of  $q$ -vectors and corresponding prices  $s$  for the customer to choose from (see e.g. Hart and Nisan, 2013), and a revenue-optimal menu must have  $s \geq 0$ , with equality when  $q$  is the zero vector. Consequently,  $\text{REV}(D)$  is an upper bound on  $\text{OPT}$ , which is only allowed to price the deterministic subsets  $q \in \{0, 1\}^n$ . We now provide a guarantee on the profit of PBDC relative to  $\text{REV}(D)$ .

**Theorem 3.2.** *On any instance with  $D$  independent, the profit of either PBDC or PC is at least  $\frac{1}{5.2} \cdot \text{REV}(D)$ .*

Several remarks are in order. First,  $\text{REV}(D)$  is a tighter upper bound on  $\text{OPT}$  than the expected welfare  $\mu_w$ , which could be  $\infty$  when  $\nu_w$  is unaccounted for. Second, the inclusion of PC in the statement of Theorem 3.2 is necessary, in that there are instances on which PB alone does not achieve a constant factor, as shown in Hart and Nisan (2017). Theorem 3.2 is a generalization of the celebrated result of Babaioff et al. (2020), which says that without costs, the profit of either PB or PC is at least  $\frac{1}{6} \cdot \text{REV}(D)$ .

The first contribution of Theorem 3.2 is that it applies to settings with production costs. Our proof first eliminates these costs by analyzing the cost-adjusted valuations  $x_i - c_i$  instead. However, the cost-adjusted valuations could be negative, which existing techniques do not handle. Furthermore, one cannot increase the negative valuations to 0 and analyze  $\max\{x_i - c_i, 0\}$  instead, because increasing the valuations could increase the optimum  $\text{REV}(D)$  against which we are comparing (see Hart and Nisan, 2017). Consequently, our analysis must show how to define concepts from Babaioff et al. (2020) such as the “marginal mechanism” for negative valuations.

The second contribution of Theorem 3.2 is that it improves the bound from  $\frac{1}{6}$  to  $\frac{1}{5.2}$ . This is obtained by analyzing the *core* and the *tail* in the decomposition of Babaioff et al. (2020) together, and showing that the worst case for PBDC in the core and worst case for PC in the tail cannot simultaneously occur.

#### 3.1 Upper Bound on the Guarantee

Finally, we present an upper bound to complement the lower bound presented in Theorem 3.2. We first show how to construct an example where *Mixed Bundling* (MB), the pricing scheme of selling items individually but offering a bundle discount for purchasing all the

items, performs much better than either PB or PC.

*Example 3.3.* Consider an instance with 2 costless items, which have IID valuations distributed as follows. There is a point mass of size  $1 - \rho$  at 0, a point mass of size  $\frac{\rho}{2}$  at 2, and the remaining  $\frac{\rho}{2}$  mass distributed in an *equal-revenue* fashion on  $[1, 2]$ , i.e. selling individually at any price in  $[1, 2]$  results in the same revenue. Formally, if  $Y$  is a random variable with this distribution, then

$$\mathbb{P}[Y \geq y] = \begin{cases} 1 & y = 0 \\ \rho & 0 < y \leq 1 \\ \frac{\rho}{y} & 1 \leq y \leq 2 \end{cases}$$

where the value of  $\rho$  is optimized to be  $\frac{3}{3+\ln 2} \approx 0.81$ .

**Theorem 3.4.** *Consider the instance in Example 3.3. The best possible PC revenue is  $2\rho$ , attained by selling individual items at any price in  $[1, 2]$ . The best possible PB revenue is also  $2\rho$ , attained by selling the bundle at the price of 2 or 3. The optimal revenue is at least  $2\rho(2 - \rho)$ ; this value can be achieved by selling individual items at the price of 2, and the bundle at the discounted price of 3.*

Therefore, neither PC nor PB, nor any mechanism that partitions<sup>3</sup> the items into bundles, can obtain more than  $\frac{3+\ln 2}{3+2\ln 2} \cdot \text{REV}(D)$  which is approximately

$$\frac{1}{1.19} \cdot \text{REV}(D).$$

Theorem 3.4 provides the best-known upper bound on the broad class of partitioning mechanisms, which captures both PC and PB. There is a tighter upper bound on the performance of just PC and PB, of  $1/2$ , due to Rubinstein (2016). However, in his example, PC and PB both perform poorly because they do not split the items into those which are best sold individually and those which are best sold together; in his example partitioning mechanisms are optimal. By contrast, in our example even partitioning mechanisms are suboptimal, because they do not “price-discriminate”, i.e. allow customers who highly value an item to buy it for its individual price, but still give customers with lower valuations a chance of buying it as part of a discounted bundle.

<sup>3</sup>A *partitioning mechanism* first splits the items into groups, and then sells each group as a bundle at the optimal bundle price for that group. Partitioning mechanisms capture PC (individual selling) because this is when each item is put into its own group. They also capture PB (pure bundling) because this is when all items are put into the same group.

## 4 Numerical Experiments

In this section we repeat the numerical experiments from Chu et al. (2008), with PBDC being added as a pricing scheme to be compared.

### 4.1 Procedure

For consistency, we follow the setup from Chu et al. (2008) as closely as possible. We use the same five families of valuation distributions commonly used to model demand—Exponential, Logit, Lognormal, Normal, and Uniform. We also use the same ranges of parameters for these families, as outlined in Table 1. The parameters were calibrated so that valuations across different families have similar means on average, and the highest means are around 10 times the lowest means. We allow for free disposal, just like Chu et al. (2008)—all negative valuations are converted to 0. We assume that valuations are independent across items.

Table 1: Ranges of Parameters (from Chu et al., 2008)

Exponential	Marginal distributions are Exponential, with means chosen uniformly from $[0.2, 2]$ . Thus the rates $\lambda$ are in $[0.5, 5]$ .
Logit	Marginal distributions are Gumbel, with fixed scale $\sigma = 0.25$ and means chosen uniformly from $[0, 2.5]$ . Thus the locations $\mu$ are in $[-0.25\gamma, 2.5 - 0.25\gamma] \approx [-0.14, 2.36]$ .
Lognormal	Marginal distributions are Lognormal. Logarithms of valuations are Normally distributed with means chosen uniformly from $[-1.5, 1]$ and fixed variance $\sigma^2 = 0.25$ . Thus the original valuations have means in $[e^{-1.5+0.125}, e^{1+0.125}] \approx [0.25, 3.08]$ .
Normal	Marginal distributions are Normal with means chosen uniformly from $[-1, 2.5]$ and variances chosen uniformly from $[0.25, 1.75]$ .
Uniform	Marginal distributions are Uniform on $[0, b]$ , where $b$ is chosen uniformly from $[0.4, 4]$ . Thus the means are in $[0.2, 2]$ .

As far as costs, we consider three scenarios:

1. *Heterogeneous Items:* in this scenario we allow valuation distributions to fluctuate in accordance to Table 1 while costs are kept low. Specifically, the cost of each item is set to 0.2, except in the case of Uniform distributions, where it is set to

half the item's mean valuation. These are the same numbers used in Chu et al. (2008).

2. *Heterogeneous Costs*: in this scenario we keep the valuation distributions identical while allowing costs to fluctuate. In the cases of the Exponential, Logit, Lognormal, or Normal distributions, the valuation is fixed to have mean 1.25, 1.5,  $e^{0.5+0.125} \approx 1.87$ , or 1.5 (with fixed variance 1) respectively; the costs are chosen uniformly from  $[0, 2.5]$ , approximately the same range as the means. In the case of the Uniform distribution, the costs are chosen uniformly from 0 to 0.75 times the maximum valuation  $b$  drawn according to Table 1. Generally in this scenario we have chosen the fixed means to lie in the middles of the ranges from Table 1.
3. *Heterogeneous Items and Costs*: in this scenario we allow both valuation distributions and costs to fluctuate (independently) as described in the preceding scenarios.

We compare four pricing schemes—PC, PB, BSP, and PBDC and consider  $n$  from 2 up to 6, which captures the range of experiments in Chu et al. (2008). For each combination of the 3 cost scenarios, 5 demand distributions, and 5 options for  $n$ , we randomly generate 200 instances, resulting in 15000 total instances. Chu et al. (2008) were able to discretize the parameter space for each combination and generate 220 instances in a grid. While generating instances in a grid is more reliable, we have too many combinations to do so, since we allow costs to vary independently. Our randomized approach also has the advantage of not depending on the exact grid of parameters chosen.

## 4.2 Observations

First we report the performance of the pricing schemes separated by scenario. For each instance (out of the 15000), we compute which of PC, PB, BSP, PBDC earns the maximum profit on that instance, and record the performance of every pricing scheme as a *fraction* of this maximum. For each scenario (out of the 3), we report the median performance as well as 10'th percentile performance of every pricing scheme across the 1000 instances of each distribution family (200 for each of  $n = 2, \dots, 6$ ), in Table 2.

We know from Chu et al. (2008) that BSP is within 1% of the optimal deterministic pricing in most of their settings, so there is minimal room for improvement under scenario 1. In fact, PBDC is a special case of BSP when all costs are identical, and very similar to PB when costs are low. However, as one can see in Table 2, PBDC still extracts close to 100% of the near-optimal

Table 2: Median and 10'th Percentile Performances

Heterogeneous Items	PC	PB	BSP	PBDC	
Exponential	0.1 %ile	.766	.940	<b>1</b>	.994
	0.5 %ile	.835	.972	<b>1</b>	.999
Logit	0.1 %ile	.826	.937	<b>1</b>	.988
	0.5 %ile	.873	.992	<b>1</b>	.998
Lognormal	0.1 %ile	.734	.982	<b>1</b>	.998
	0.5 %ile	.799	.996	<b>1</b>	<b>1</b>
Normal	0.1 %ile	.825	.745	<b>1</b>	.957
	0.5 %ile	.890	.880	<b>1</b>	.975
Uniform	0.1 %ile	.904	.834	.940	<b>.949</b>
	0.5 %ile	.959	.867	.975	<b>.998</b>
Heterogeneous Costs	PC	PB	BSP	PBDC	
Exponential	0.1 %ile	.850	.269	.807	<b>.995</b>
	0.5 %ile	.931	.489	.907	<b>1</b>
Logit	0.1 %ile	.815	.063	.245	<b>.996</b>
	0.5 %ile	.891	.481	.595	<b>1</b>
Lognormal	0.1 %ile	.775	.513	.760	<b>1</b>
	0.5 %ile	.861	.730	.880	<b>1</b>
Normal	0.1 %ile	.858	.297	.779	<b>.982</b>
	0.5 %ile	.926	.547	.912	<b>1</b>
Uniform	0.1 %ile	.872	.348	.875	<b>.948</b>
	0.5 %ile	.933	.578	.974	<b>1</b>
Both Heterogeneous	PC	PB	BSP	PBDC	
Exponential	0.1 %ile	.884	.137	.759	<b>.978</b>
	0.5 %ile	.964	.403	.926	<b>1</b>
Logit	0.1 %ile	.852	.001	.385	<b>.987</b>
	0.5 %ile	.938	.168	.894	<b>1</b>
Lognormal	0.1 %ile	.852	.015	.327	<b>.931</b>
	0.5 %ile	.970	.245	.887	<b>1</b>
Normal	0.1 %ile	.904	.010	.699	<b>.974</b>
	0.5 %ile	.978	.198	.933	<b>1</b>
Uniform	0.1 %ile	.914	.380	.605	<b>.937</b>
	0.5 %ile	.982	.638	.875	<b>1</b>

BSP profit under this scenario. For Uniform valuations, PBDC is no longer a special case of BSP, since costs vary proportionally with means. PBDC actually outperforms BSP in this setting—indeed, this is by far the worst setting for BSP listed in Chu et al. (2008, tbl. 5), where it only extracts 91% of the optimum.

Scenario 2, in which valuation distributions are identical but costs are allowed to fluctuate, really exhibits the power of PBDC, which allows customers to consume only the items they value above cost via self-selection. PC loses out on not bundling similar items that differ only in cost, while BSP is forced to compromise between charging cheap prices which cause overinclusion loss in the high-cost items, or charging expensive prices which cause deadweight loss in the low-cost items.<sup>4</sup> In the appendix of the full version of this pa-

<sup>4</sup>The welfare (which is fixed) can be decomposed as the sum of the firm's profit ("producer surplus"), the customer's utility minus cost ("consumer surplus"), the sum of  $x_i - c_i$  over items not sold which should have been sold because  $x_i > c_i$  (this is the "deadweight loss"), and the sum of  $c_i - x_i$  over items sold which should not have been sold because  $c_i > x_i$  (this is the "overinclusion loss"). In gen-

per (Ma and Simchi-Levi, 2015), we show an instance that exemplifies why BSP performs poorly when the costs in the setup from Chu et al. (2008) are increased.

When both valuation distributions and costs are allowed to vary under scenario 3, PBDC is still the best strategy by a significant margin. However, the benefits of bundling have decreased when items can be drastically different, and consequently PC has gained ground. It seems intuitive to hypothesize that the performance of PC is inflated by the small values of  $n$  we are using. In the next subsection, we organize our reports separated by  $n$ , under scenario 3 where both valuation distributions and costs fluctuate.

### 4.3 Separation by $n$ and Breakdown of Welfare

In this subsection, we allow both valuation distributions and costs to fluctuate, and report averages across demand distributions, separated by  $n$ . Since the distribution families we’re amalgamating were calibrated to have similar means over their ranges of parameters once  $n$  is fixed, it makes sense in this subsection to report average absolute profits, instead of median fractions. We also report the “economics figures” defined in Footnote 4, in the spirit of Chu et al. (2008), in Table 3. The main conclusions are then summarized in graphs.

The first graph (Figure 1) shows that although PBDC optimizes from the perspective of a selfish monopolist interested only in profit, it has a similar advantage in terms of sum of producer and consumer surplus. Indeed, there is zero overinclusion loss, and the monopolist is encouraged to choose a bundle price low enough to accommodate most customers. PC also incurs no overinclusion loss, but incurs more deadweight loss because it does not bundle. PB incurs significantly more overinclusion loss than any other strategy, forcing the customer into buying every item at once. All in all, PBDC is equally attractive from the long-term perspective of maximizing the customers’ surplus.

The second graph (Figure 2) shows the profits of each pricing scheme as  $n$  increases. PC profits increase linearly with  $n$ , since items are sold separately. Both PB and BSP profits are concave in  $n$ —that is, the marginal gain from having one more item to sell is decreasing. Indeed, PB is burdened with adding to its grand bundle another item that could be valued below cost, while BSP is burdened with an additional distinct item to consider in its item-symmetric price structure. PBDC is the only pricing scheme where the profit curve is (slightly) convex in  $n$ , since each

general, to increase profit and consumer surplus one aims to minimize deadweight loss and overinclusion loss.

Table 3: Report of Economics Figures

$n$	Statistic	PC	PB	BSP	PBDC
2	Producer Surplus	0.427	0.301	0.412	<b>0.432</b>
	Consumer Surplus	0.287	0.194	0.250	0.292
	Total Surplus	0.714	0.495	0.662	0.724
	Deadweight Loss	0.192	0.351	0.224	0.183
	Overinclusion Loss	-	0.061	0.021	-
3	Producer Surplus	0.655	0.395	0.630	<b>0.683</b>
	Consumer Surplus	0.437	0.254	0.382	0.436
	Total Surplus	1.092	0.649	1.011	1.119
	Deadweight Loss	0.291	0.604	0.352	0.264
	Overinclusion Loss	-	0.130	0.020	-
4	Producer Surplus	0.870	0.457	0.827	<b>0.929</b>
	Consumer Surplus	0.587	0.293	0.497	0.582
	Total Surplus	1.456	0.749	1.324	1.511
	Deadweight Loss	0.396	0.905	0.498	0.342
	Overinclusion Loss	-	0.198	0.031	-
5	Producer Surplus	1.070	0.504	1.030	<b>1.167</b>
	Consumer Surplus	0.705	0.297	0.595	0.703
	Total Surplus	1.775	0.802	1.625	1.870
	Deadweight Loss	0.488	1.158	0.600	0.394
	Overinclusion Loss	-	0.304	0.039	-
6	Producer Surplus	1.265	0.553	1.206	<b>1.409</b>
	Consumer Surplus	0.844	0.346	0.697	0.828
	Total Surplus	2.108	0.899	1.902	2.237
	Deadweight Loss	0.587	1.440	0.736	0.459
	Overinclusion Loss	-	0.356	0.057	-

item creates additional incentive for the customer to purchase the bundle, and makes the customer’s total utility from purchasing more concentrated about its mean. This confirms the hypothesis that while Table 2 reports a small gap between PC and PBDC under scenario 3, this gap quickly widens as  $n$  increases.

To summarize our numerical experiments, we considered both scenarios with low costs and scenarios with high costs, and reported median performances over  $n = 2, \dots, 6$  for different demand distributions. When costs are low, PC can earn as little as 79.9% of the profit of the best mechanism among PC, PB, BSP, and PBDC. When costs are high, PB can earn as little as 16.8% of the profit of the best mechanism, BSP can earn as little as 59.5%, and PC also falls behind as  $n$  increases. PBDC has the highest percentages overall, and is by far the most robust over different cost scenarios, always obtaining at least 97.5% of the profit of the best mechanism among PC, PB, BSP, and PBDC. We should point out that throughout our simulations, PBDC was also computationally much faster than BSP, requiring an optimization over 1 parameter instead of  $n$  parameters.

### Acknowledgements

The authors would like to thank anonymous reviewers as well as Aviad Rubinstein whose excellent suggestions helped improve the paper.

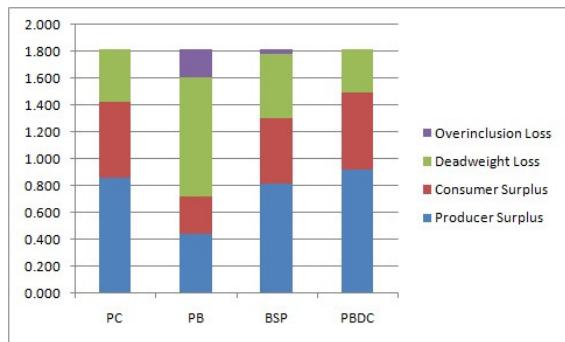


Figure 1: Breakdowns of Welfare, averaged over  $n$

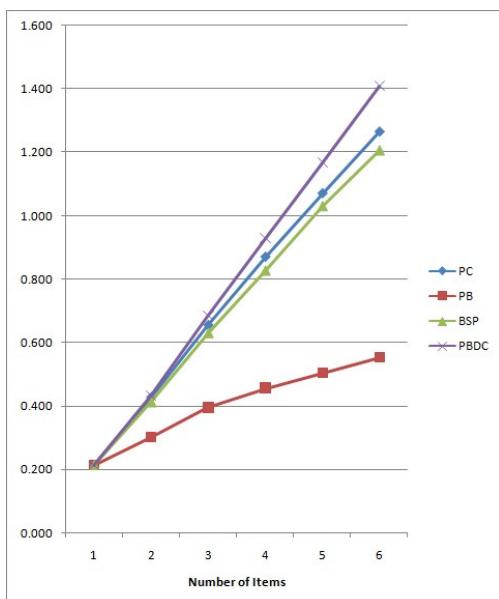


Figure 2: Average Profits, as a function of  $n$

## References

W. J. Adams and J. L. Yellen. Commodity bundling and the burden of monopoly. *The quarterly journal of economics*, pages 475–498, 1976.

M. Armstrong. Price discrimination by a many-product firm. *The Review of Economic Studies*, 66(1):151–168, 1999.

M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A simple and approximately optimal mechanism for an additive buyer. *Journal of the ACM (JACM)*, 67(4):1–40, 2020.

Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits, and efficiency. *Management science*, 45(12):1613–1630, 1999.

M. Bateni, S. Dehghani, M. Hajiaghayi, and S. Seddighin. Revenue maximization for selling multiple correlated items. In *Algorithms-ESA 2015*, pages 95–105. Springer, 2015.

Y. Cai, N. R. Devanur, K. Goldner, and R. P. McAfee. Simple and approximately optimal pricing for proportional complementarities. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 239–240, 2019.

H. Chen, M. Hu, and G. Perakis. Distribution-free pricing. *Available at SSRN 3090002*, 2019.

C. S. Chu, P. Leslie, and A. Sorensen. Nearly optimal pricing for multiproduct firms. Technical report, National Bureau of Economic Research, 2008.

C. S. Chu, P. Leslie, and A. Sorensen. Bundle-size pricing as an approximation to mixed bundling. *The American Economic Review*, pages 263–303, 2011.

R. Ernst and P. Kouvelis. The effects of selling packaged goods on inventory decisions. *Management Science*, 45(8):1142–1155, 1999.

H. Fang and P. Norman. To bundle or not to bundle. *RAND Journal of Economics*, pages 946–963, 2006.

S. Hart and N. Nisan. *The menu-size complexity of auctions*. Center for the Study of Rationality, 2013.

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

S. Hart and P. J. Reny. Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoretical Economics*, 10(3):893–922, 2015.

L. M. Hitt and P.-y. Chen. Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Science*, 51(10):1481–1493, 2005.

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

X. Li, H. Sun, and C. P. Teo. Convex optimization for bundle size pricing problem. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 637–638, 2020.

G. Lugosi. *Concentration-of-measure inequalities*. <http://www.econ.upf.edu/~lugosi/anu.pdf>, 2009.

W. Ma and D. Simchi-Levi. Reaping the benefits of bundling under high production costs. *arXiv preprint arXiv:1512.02300*, 2015.

R. P. McAfee, J. McMillan, and M. D. Whinston. Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics*, pages 371–383, 1989.

A. Rubinstein. On the computational complexity of optimal simple mechanisms. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 21–28. ACM, 2016.

A. Rubinstein and S. M. Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. *ACM Transactions on Economics and Computation (TEAC)*, 6(3-4):1–25, 2018.

R. Schmalensee. Gaussian demand and commodity bundling. *Journal of business*, pages S211–S230, 1984.

G. J. Stigler. United states v. loew's inc.: A note on block-booking. *Sup. Ct. Rev.*, page 152, 1963.

R. Venkatesh and V. Mahajan. 11 the design and pricing of bundles: a review of normative guidelines and practical approaches. *Handbook of pricing research in marketing*, page 232, 2009.

R. B. Wilson. *Nonlinear pricing*. Oxford University Press, 1993.

A. C.-C. Yao. 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, 2015.

Y. Zhao. *Inequalities*. <http://yufeizhao.com/wc08/ineq.pdf>, 2008.