Bundling Sequentially Released Durable Goods

Luís Cabral
New York University

Gabriel Natividad
Universidad de Piura

February 2021

Abstract. Suppose two durables are sequentially released and suppose that consumer valuations of these goods are positively correlated. By the time the second good is released, high-valuation buyers are out of the market for the first good (for they bought it when first released). Therefore, a bundle can be targeted at the low-valuation consumers without violating the high-valuation consumers' incentive compatibility constraint. We test the model's predictions on data from retail DVD sales in the 2000s. Consistent with theory, our estimates suggest that mixed bundling increases revenues, especially when the bundle components are similar (which in turn suggests positive correlation of valuations).

Cabral: Stern School of Business, New York University; luis.cabral@nyu.edu. Natividad: Universidad de Piura; gabriel.natividad@udep.edu.pe. We are grateful to the Editor, three referees, and seminar participants at NYU, Harvard, Columbia, Georgia Tech, North Carolina, Toronto, and the 2017 Munich Summer Institute for useful comments, and Dainis Zegners for a discussion. Cristián Figueroa, Diego Zúñiga and Timothy Petliar provided valuable research assistance. The usual disclaimer applies.

1. Introduction

At least since Stigler (1963), the practice of bundling has been viewed as a form of second-degree price discrimination that takes advantage of the negative correlation in buyer valuations. Surveying this literature, Chen and Riordan (2013) argue that

> A multiproduct monopolist generally achieves higher profit from mixed bundling than from separate selling if consumer values for two of its products are negatively dependent, are independent, or have sufficiently limited positive dependence.

To understand the standard argument for bundling, suppose that a seller offers two products, x and y, and that buyers have valuation $\overline{u}$ or $\underline{u}$ for each of these products, with $\overline{u} > \underline{u}$. If valuations are negatively correlated (a consumer with high valuation of x has low valuation of y), then (pure) bundling allows the seller to increase profits: bundling "homogenizes" demand, so that by setting the bundle price at $\overline{u} + \underline{u}$, the seller is able to sell to all buyers and extract all consumer surplus. By contrast, if valuations are positively correlated (a consumer with high valuation of x also has a high valuation of y), then bundling does not help the seller: either the seller offers a high price that is accepted by high-valuation buyers or the seller offers a low price that is accepted by all buyers, just as in the no-bundling case.

The calculus of bundling changes dramatically if we consider sequentially released durable goods. Based on numerical simulations, Derdenger and Kumar (2013) show that bundling may provide a means for dynamic consumer segmentation. Specifically, bundles "attract some segments of consumers to advance their purchases and others to enter the market when they might not have otherwise" (p. 853). In the present study, we follow this line of argument by developing and testing a theoretical model of dynamic price discrimination with durable goods. Suppose that good x is a durable good sold over two periods, $t = 1$ and $t = 2$, whereas good y is only sold at $t = 2$. If high-valuation buyers make a purchase at $t = 1$, then, when $t = 2$ comes around, high-valuation buyers are effectively out of the market for x. This is true in the standard durable-goods monopoly model, and is the basis for the strategy of price-skimming: once high-valuation buyers have left the market, the seller can lower price and capture additional value from low-valuation buyers. The element we add is that, to the extent that valuations of x and y are positively correlated, there may be scope for profit-increasing bundling at $t = 2$. The idea is simple: absent bundling, the seller optimally targets the price of y at high-valuation buyers only. However, by offering a bundle of x and y targeted at low-valuation buyers, the seller is able to increase the sales of y without lowering the price paid by high-valuation buyers.

To fix ideas, consider the videogame industry. During the early 2000s Nintendo was a de facto monopolist in the portable videogame console market. The Nintendo GBA console was first released in June 2001. In November 2002, Nintendo started selling a bundle of the Nintendo GBA console together with the game Mario Advance 2. The game could also be purchased separately. More generally, the videogame industry provides multiple examples of mixed bundling of hardware and software. At first, it may seem hardware/software bundling is not a good example because product complementarity might be the reason for bundling. However, while the value of two products consumed jointly is greater than the sum of their separate values, the value of the set is the same regardless of whether they were bought separately or as a bundle. In other words, if a bundle is a pricing deal, not a separate product per se, then the above reasoning regarding bundling continues to apply.

We derive a set of conditions such that bundling is optimal in the sequential release case even though it is not optimal in the one-period case. Our basic result considers the extreme setting when valuations are perfectly correlated. We then consider the extension to positively but not perfectly correlated valuations and show that the gain from bundling is greater the higher the correlation of valuations. In a sense, durability and sequential release turn the calculus of the benefits from bundling on its head: in the static framework, bundling plays no role when valuations are perfectly correlated. By contrast, in the dynamic context bundling plays a role precisely because valuations are perfectly (or close to perfectly) correlated.

We then test the above theoretical predictions on data from the DVD retail industry in the U.S. from 2000 to 2009. Many DVD titles are sold in retail stores in bundles, typically as a bundle of two different titles. We document the extent to which bundles are biased towards selecting similar titles (they are). This suggests our theoretical model of positive correlations in valuations applies. By means of a simple regression analysis (with multiple controls), we provide an estimate of the gains from mixed bundling. The estimates we obtain are rather large — between 30 and 40% — and statistically precise. Moreover, we estimate that the gain from mixed bundling is greater the greater the similarity between bundled titles.

■ **Related literature.** The paper that is closest to ours is Derdenger and Kumar (2013). They structurally estimate a model of demand for hardware (videogame consoles) and software (videogames). By means of numerical counterfactuals, they show that bundling software with hardware may improve a strategy of intertemporal price discrimination. Moreover, they show that an increase in correlation of valuations leads to an increase in revenues from bundling. Our Proposition 1 and Corollary 1 corroborate these results in Derdenger and Kumar (2013). In fact, our theoretical model is a distinguishing feature with respect to their work, as Derdenger and Kumar (2013) offer no theoretical foundation for the strategy of combining dynamic pricing and bundling. We believe formal modeling offers two key advantages to address the question of interest here. First, it helps sharpen assumptions, results and intuitions. For example, Derdenger and Kumar (2013) "find that bundles serve a role similar to an additional product in the firm's product offering because consumers do not value the bundle identical to the sum of valuations of the component products" (p. 828). However, we show that the mechanism of dynamic market segmentation does not depend on there being any complementarity between bundle components (and the empirical tests we consider arguably feature no complementarities). Second, finding a formal equilibrium gives some breadth to the argument. Absent a formal model of seller behavior, Derdenger and Kumar (2013) approximate the seller's pricing decisions by an AR1 process, which may limit the strength of the exercise.

Our paper relates to two theories of price discrimination: bundling and dynamic durable-good pricing. As mentioned earlier, since Stigler's (1963) seminal paper a large literature has derived the conditions for revenue increasing bundling as a function of the distribution of buyer preferences (e.g., Adams and Yellen, 1976, McAfee et al., 1989). However, the approach is typically static, that is, it assumes one-time pricing and purchase decisions. Regarding dynamic pricing of durable goods, the seminal paper by Coase (1972) gave rise to an extensive literature looking primarily at dynamic pricing as a form of price discrimination and at the implications of commitment — or lack thereof — to future prices (e.g., Stokey, 1981; Bulow, 1982). To the best of our knowledge, this literature does not address the issue of bundling. Our contribution is primarily to bring these two strands of the literature together: we look at price discrimination when the seller combines dynamic pricing and bundling.

We are not the first to suggest that bundling may be a profitable strategy when valuations are positively correlated. In addition to Derdenger and Kumar (2013), Gandal et al. (2018) show that there are gains from bundling in a static framework even when valuations are positively correlated. The idea is that the Stigler (1963) model misses an important point, which Gandal et al. (2018) refer to as the market-expansion effect of bundling. Specifically, consider the case of *pure* bundling and assume that the share of actual purchasers is a small fraction of the population. If valuations are positively correlated across products, then the effect of bundling is to "fatten" the tail of the distribution of valuations (variance-increasing effect). As Johnson and Myatt (2006) show (theoretically) and Gandal et al. (2018) observe (empirically), this is consistent with a revenue-increasing effect of bundling. This is a very different point from ours. In fact, it depends on the seller's strategy being one of pure bundling (or approximately pure bundling), which is not the case in our theoretical model or empirical tests.

We are not aware of many other economics papers that estimate the effects of bundling empirically (in addition to Derdenger and Kumar, 2013, and Gandal et al., 2018). Crawford and Yurukoglu (2012) estimate a structural model of cable TV demand and run a series of unbundling counterfactuals. They show that the total and consumer welfare impact varies across agents (that is, some suppliers win, some lose, and some consumers win while others lose). Other empirical papers that analyze bundling include Gentzkow (2007), who studies joint purchases of print and online newspapers, Chu et al. (2011), who estimate the demand for bundled theater tickets, and Ho et al. (2012), who analyze welfare effects of full-line forcing in the video rental industry.

2. Theory

In this section we present our theoretical framework as well as our main results. We first consider the case when there are no dynamics. This reference point is important for, as we will show, within the confine of our model there is no scope for profit-increasing bundling. By contrast, when we extend the model to sequential releases we find that, under some parameter assumptions, bundling strictly increases seller profit. Although for most of the section we consider the extreme case of perfectly correlated valuations, we end the section with an extension to imperfectly correlated valuations.

■ **Basic model.** Consider a seller with two goods, x and y, produced at zero marginal cost. There is a measure one of buyers who are willing to purchase at most one unit of each good. Buyer valuation can either be high, $\overline{u}$, or low, $\underline{u}$, with $\overline{u} > \underline{u} > 0$; and a fraction α of buyers have high valuation.

■ **Static case.** Although our contribution to the bundling literature is on the effect of time in bundling durable goods, it helps to consider the static case as a reference point. Suppose there is only one period and that the two products are available. Absent bundling, optimal pricing can be studied in isolation. Consider for example the optimal price of x. The only candidates for optimal price are $p_x = \overline{u}$ and $p_x = \underline{u}$. The first one leads to profit $\pi = \alpha\,\overline{u}$, the second one to $\pi = \underline{u}$. It follows that if $\alpha > \underline{u}/\overline{u}$ then the seller is better off by setting $p_x = \overline{u}$, whereas if $\alpha < \underline{u}/\overline{u}$ then the seller is better off by setting $p_x = \underline{u}$.

Consider now the possibility of offering a bundle of x and y for a price of p_{xy}. At this point, it is crucial to know how the valuations of x and y are correlated (if at all). Specifically, suppose that the valuations of x and y are (perfectly) negatively correlated. In order to maintain symmetry, suppose that $\alpha = \frac{1}{2}$, so that half the consumers have high valuation of x and low valuation of y, whereas the other half have high valuation of y and low valuation of x. Then by setting $p_{xy} = \overline{u} + \underline{u}$ the seller makes a total profit of $\overline{u} + \underline{u}$. By contrast, by selling x and y as separate items, the seller's maximum profit is given by $\max\{\overline{u}, 2\,\underline{u}\}$, which is strictly lower than $\overline{u} + \underline{u}$. This corresponds to the old intuition in Stigler (1963), aptly summarized in Chen and Riordan (2013): "A multiproduct monopolist generally achieves higher profit from mixed bundling than from separate selling if consumer values for two of its products are negatively dependent."

Consider now the opposite extreme, that is, the case when valuations are perfectly positively correlated: a fraction α has valuation $\overline{u}$ for both x and y, whereas a fraction $(1 - \alpha)$ has valuation $\underline{u}$ for both x and y. Consistent with the Stigler (1963) and Chen and Riordan (2013) perspective, we conclude that in this case bundling does not provide any increase in seller profit. If $\alpha > \underline{u}/\overline{u}$ then the seller is better off by setting $p_{xy} = 2\,\overline{u}$, whereas if $\alpha < \underline{u}/\overline{u}$ then the seller is better off by setting $p_{xy} = 2\,\underline{u}$. However, this pricing strategy implies exactly the same profit as optimal pricing of individual items.

In most of what follows, we will continue with the case of perfect correlations. To the extent that we do find scope for profit increasing bundling, we will highlight the importance of dynamics in the shaping of bundling incentives.

■ **Complements.** Some of the examples where bundling takes place involve complement products, namely situations when the value of jointly consuming x and y is greater than the sum of the separate consumption values of x and y. How does this affect the above discussion on bundling? Suppose that joint consumption of x and y leads to an additional value v with respect to the sum of the separate values of x and y. The crucial question is whether this value v can be obtained by simply consuming x and y separately, or rather requires the purchase of a bundle of x and y. This is largely an empirical question. For example, in the DVD example we examine later most of the bundles consist of two DVDs put together. Even if there is an extra value from owning the two DVDs over and beyond the sum of the individual values, the buyer can attain v by purchasing x and y separately or as a bundle. By contrast, if the bundle corresponds to a "special edition" of a series of sequels, then the packaging and/or the liner notes may provide a v that is not attainable by purchasing x and y separately.

We believe many if not most of the examples of complements fall into the category where v does not require the purchase of the two units as a bundle. If that is the case, then the above discussion applies equally well to the additive valuation and to the complement case. In particular, if the valuations of x and y are positively correlated, then there is no scope for the seller to increase profits by means of bundling.

■ **Sequential release.** We now come to the core of our model. We now assume that two products, x and y, are sequentially released. Specifically, x is released at $t = 1$ and y at

$t = 2$. This means that x can be purchased at $t = 1$ or $t = 2$, whereas y can only be purchased at $t = 2$. We denote by p_{xt} product x's price at time t, and by p_y product y's price (at time $t = 2$). Finally, we also consider the possibility of selling the bundle xy at time $t = 2$ and denote the bundle price by p_{xy}. The seller discounts period 2 according to a discount factor δ_F, whereas the buyers' discount factor is given by δ_C. We assume that:

Assumption 1. $\delta_F > \delta_C < \frac{1}{2}$.

The assumption that buyers discount the future more than the seller can be justified in various ways. In particular, we should think of the discount factor as including, in addition to the time preference for money, the ability to predict future product-relevant events, and it makes sense to assume that, at $t = 1$, the seller is better informed about $t = 2$ product releases. The assumption that $\delta_F > \delta_C$ is required for there to be an equilibrium where purchases of x are spread over time. If $\delta_F = \delta_C$, then the p_{x1} values at which high-valuation buyers would be willing to purchase at $t = 1$ are too low for the seller to be willing to sell at $t = 1$. The additional constraint that $\delta_C < \frac{1}{2}$ is required, again, for the x sales to be spread over time. If $\delta_C > \frac{1}{2}$, then a high-valuation buyer is relatively more willing to wait for a lower price at $t = 2$. As a result, the seller would be forced to set p_{x1} so low that low-valuation buyers would prefer to purchase at $t = 1$ rather than wait for $t = 2$.[1]

Assumption 2. $\alpha > \underline{u}/\overline{u} > \frac{1}{2}$.

This assumption places bounds on the relative valuations. Valuations must be sufficiently close that the seller is interested in selling to both types. If $\underline{u}/\overline{u}$ were very low, then the seller would optimally ignore low-valuation buyers and the problem would become trivial. At the oppose end, if $\underline{u}/\overline{u}$ is very high, then the buyers' incentive compatibility constraints become too tight for price discrimination to be profitable.

As in the static case, we consider the case when buyer valuations are perfectly correlated: a fraction α of buyers has high valuation of both products, whereas a fraction $1 - \alpha$ has low valuation of both products. The equilibrium concept we consider is that of subgame-perfect equilibrium. In the present context, backward induction is equivalent to subgame perfection, so effectively we derive the Nash equilibrium obtained by solving the game backward, beginning with $t = 2$.

■ **The no-bundling case.** Suppose first that the seller does not resort to bundling. Consider first the price of y. Since the two goods are independent, we have a simple monopoly pricing problem. Since $\underline{u} < \alpha\,\overline{u}$ (per Assumption 2), it is optimal to set $p_y = \overline{u}$. Regarding the price of x, we have a classical durable-goods problem. If the seller cares enough about the future, the subgame-perfect equilibrium has $p_{x2} = \underline{u}$ and p_{x1} such that a high-valuation buyer is indifferent between purchasing at $t = 1$ and waiting until $t = 2$. This leads to

$$\overline{u} - p_{x1} = \delta_C\,(\overline{u} - p_{x2})$$

1. A referee rightly points out that we must also assume that δ_F is not too close to 1. In the class of models with two types and two periods like the one we consider, the unique subgame-perfect equilibrium when δ_F is close to 1 involves the seller maintaining a high price for the good in both periods ($p_{x1} = p_{x2} = \overline{u}$). The high-valuation buyers mix over purchasing in each period such that the seller is indifferent between $p_{x2} = \overline{u}$ and $p_{x2} = \underline{u}$.

or simply

$$p_{x1} = (1-\delta_C)\,\overline{u} + \delta_C\,\underline{u} \tag{1}$$

given that $p_{x2} = \underline{u}$. In order for this to be optimal for the seller, it must be that the seller is better off by selling only to high-valuation buyers in the first period. This requires that

$$\alpha\,p_{x1} + \delta_F\,(1-\alpha)\,p_{x2} > \underline{u}.$$

Substituting for the values of p_{xt}, we get

$$\alpha\left((1-\delta_C)\,\overline{u} + \delta_C\,\underline{u}\right) + \delta_F\,(1-\alpha)\,\underline{u} > \underline{u}.$$

Since $\delta_F > \delta_C$ (Assumption 1), a sufficient condition is given by

$$\alpha\left((1-\delta_C)\,\overline{u} + \delta_C\,\underline{u}\right) + \delta_C\,(1-\alpha)\,\underline{u} > \underline{u}.$$

which can be re-arranged as

$$(\alpha\,\overline{u} - \underline{u})\,(1-\delta_C) > 0$$

which is implied by Assumption 2. Seller profit under no bundling is given by

$$\pi_N = \alpha\,p_{x1} + \delta_F\left((1-\alpha)\,p_{x2} + p_y\right)$$

which, given the above values of p_{xt} and p_y, becomes

$$\pi_N = \alpha\left((1-\delta_C)\,\overline{u} + \delta_C\,\underline{u}\right) + \delta_F\left((1-\alpha)\,\underline{u} + \alpha\,\overline{u}\right). \tag{2}$$

■ **The bundling case.** Suppose now that the seller offers a bundle at $t=2$. As in the previous case, suppose that high-valuation buyers purchase x at $t=1$. This implies that *at $t=2$ high-valuation buyers have no valuation of an additional unit of x.* Consider the strategy of targeting the bundle to low-valuation buyers. Then $p_{xy} = 2\,\underline{u}$ and $p_y = \overline{u}$. The value of p_{x1} that makes high-valuation buyers indifferent between buying at $t=1$ and waiting is such that

$$\overline{u} - p_{x1} + \delta_C\,(\overline{u} - p_y) = \delta_C\,(\overline{u} + \overline{u} - p_{xy})$$

or simply

$$p_{x1} = (1-2\,\delta_C)\,\overline{u} + 2\,\delta_C\,\underline{u}$$

where we substitute $p_{xy} = 2\,\underline{u}$ and $p_y = \overline{u}$. We must add the constraint that this price is greater than $\underline{u}$, else a low-valuation buyer would want to anticipate its x purchase to $t=1$:

$$(1-2\,\delta_C)\,\overline{u} + 2\,\delta_C\,\underline{u} > \underline{u}$$

which requires $\delta_C < \frac{1}{2}$, which is true by the second part of Assumption 1. Seller profit from this bundling strategy is given by

$$\pi_B = \alpha\big((1-2\,\delta_C)\,\overline{u} + 2\,\delta_C\,\underline{u}\big) + \delta_F\left(\alpha\,\overline{u} + 2\,(1-\alpha)\,\underline{u}\right). \tag{3}$$

■ **Main result.** As mentioned earlier, if valuations are perfectly positively correlated and if both goods are offered at the same time, then there is no scope for profit-increasing bundling strategy. Our main result is that, if x and y are sequentially released, then there is an additional avenue for profit-increasing bundling, even if buyer valuations of x and y are perfectly positively correlated.

Figure 1
Distribution of valuations with imperfectly correlated valuations

		valuations of y		
		H	L	Total
valuations of x	H	$\alpha\big(\alpha\,(1-\rho)+\rho\big)$	$\alpha\,(1-\alpha)\,(1-\rho)$	α
	L	$\alpha\,(1-\alpha)\,(1-\rho)$	$(1-\alpha)\,\big((1-\alpha)\,(1-\rho)+\rho\big)$	$1-\alpha$
	Total	α	$1-\alpha$	1

Proposition 1. *In equilibrium, the seller is strictly better off by offering a bundle at $t=2$.*

Proof: The difference between profit with bundling, (3), and profit with no bundling, (2), is given by

$$\begin{aligned}
\pi_B - \pi_N &= \alpha\big((1-2\,\delta_C)\,\overline{u} + 2\,\delta_C\,\underline{u}\big) + \delta_F\,\big(\alpha\,\overline{u} + 2\,(1-\alpha)\,\underline{u}\big) - \\
&\qquad -\Big(\alpha\,\big((1-\delta_C)\,\overline{u} + \delta_C\,\underline{u}\big) + \delta_F\,\big((1-\alpha)\,\underline{u} + \alpha\,\overline{u}\big)\Big) \\
&= \alpha\,\delta_F\overline{u} - (\overline{u}-\underline{u})\,\big((1-\alpha)\,\delta_F + \alpha\,\delta_C\big) \\
&> \alpha\,\delta_F\overline{u} - (\overline{u}-\underline{u})\,\big((1-\alpha)\,\delta_F + \alpha\,\delta_F\big) \\
&= \big(\alpha\,\overline{u} - (\overline{u}-\underline{u})\big)\;\delta_F \\
&> \big(\underline{u} - (\overline{u}-\underline{u})\big)\;\delta_F \\
&> (\underline{u}-\underline{u})\;\delta_F \\
&= 0.
\end{aligned}$$

In the above sequence, the first inequality follows from Assumption 1, namely the assumption that $\delta_F > \delta_C$; the second inequality follows from Assumption 2, namely the assumption that $\underline{u} < \alpha\,\overline{u}$; and the third inequality follows from Assumption 2, namely the assumption that $\underline{u}/\overline{u} > \frac{1}{2}$. In fact, this is equivalent to $2\,\underline{u} > \overline{u}$, or simply $\underline{u} > \overline{u} - \underline{u}$. ■

■ **Imperfect correlation.** So far we have made the rather extreme assumption that valuations are perfectly correlated. We now consider the case of imperfect correlation. The goal is two-fold: first, to show that Proposition 1 is not a knife-edged result, that is, it does not depend on the extreme assumption of perfect correlation; and second, to evaluate the relation between the degree of correlation and the seller's gain from implementing a bundling strategy.

There are many joint distributions of valuations of (x, y) that are consistent with the marginal distributions considered before, namely a fraction α with high valuation and a fraction $1-\alpha$ with low valuation. Figure 1 depicts one possible parameterization of joint valuations. For example, the cell (H, H) indicates that a fraction $\alpha\,\big(\alpha\,(1-\rho)+\rho\big)$ of all consumers have a high valuation of x and for y.

This parameterization has the advantage of (a) being consistent with the marginal distributions considered before, and (b) depending on one single parameter, ρ, which measures the degree of correlation in valuations. The value of ρ is not equal to the Pearson coefficient of correlation. However, the perfect-correlation case we considered before corresponds to

$\rho = 1$, whereas $\rho = 0$ implies independent valuations. Moreover, the coefficient of correlation is monotonic in ρ.

Proposition 1 refers to the case when $\rho = 1$. The next result corresponds to the case when ρ falls in the neighborhood of 1.

Corollary 1. *In the neighborhood of $\rho = 1$, the seller's gain from bundling is strictly increasing in ρ.*

Proof: First, seller profit remains the same under no bundling. In fact, under no bundling only the marginal distributions of valuations matter, and these are constant with respect to ρ. Second, Proposition 1 is based on strict inequalities, that is, the optimal solution is strictly better than the alternative. This implies that, if ρ is close to 1, then it remains as an optimal solution. Finally, it is straightforward to check that the revenue loss is decreasing in ρ. ■

■ **Negatively correlated valuations.** The value added by our theory model is to consider the role played by bundling when valuations are positively correlated. That said, one may also inquire into the role of dynamic price discrimination and bundling when valuations are negatively correlated. Consider the opposite extreme of the case we considered above. Suppose that there are two types of buyers: one half with valuations $\overline{u}, \underline{u}$ for x and y (type 1), respectively, and one half with valuations $\underline{u}, \overline{u}$ for x and y, respectively (type 2). We can show that, if δ_F is sufficiently small, then there is no room for profitable bundling. Pricing proceeds according to the "typical" pattern of pricing durables: type 1 consumers purchase x at $t = 1$ for $p \approx \overline{u}$. At $t = 2$, type 2 purchase y for $\overline{u}$ and x for $\underline{u}$, whereas type 1 make no purchase.

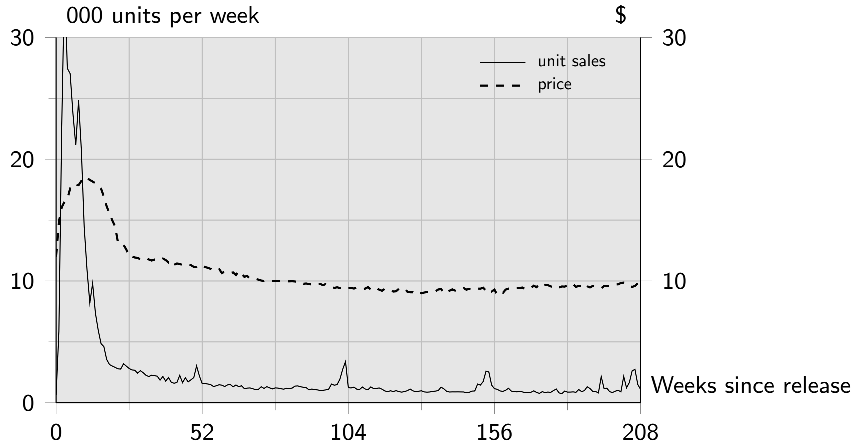
At the opposite extreme, if δ_F is sufficiently high, then there is no room for dynamic pricing: all sales take place at $t = 2$. We thus have essentially a static problem: the seller offers a bundle of $\overline{u} + \underline{u}$ which both types purchase. This is the maximum the seller can get and can only be gotten by selling a bundle.

3. Empirical evidence: data

In this section and the next we provide evidence on Propositions 1 and Corollary 1 from Section 2. The setting for our empirical application is the U.S. home DVD sales industry during the period 2000–2009.[2] In essence, the DVD sales industry comprises two stages in the value chain: content distribution companies, such as Warner Bros., selling video titles to retail channels such as Kmart, who then sell them to the final consumer.[3] While distributors are large and in small number, retailers range from fairly small specialty stores to larger retail outlets such as Amazon.com.[4]

2. A brief description of this industry is provided by Elberse and Oberholzer-Gee (2007). In many ways, the industry we study resembles the video rental industry, which has been studied extensively by Mortimer (2008). However, there are also important differences, both in the nature of demand and in the structure of the value chain.
3. Cabral and Natividad (2016) focuses on the wholesale segment of the industry. By contrast, this and the next section focus on the retail segment.
4. Upstream, distributors obtain content from a series of industries such as feature film, TV and cable producers.

Figure 2
Median sales (units) and price over time



■ **Data and summary statistics.** We use proprietary data from Nielsen VideoScan, a leading provider of information on video sales. VideoScan covers a large sample of retail outlets (but not Wal-Mart). It details weekly U.S. units sold of each video title on 24,451 feature films with active sales between 2000 and 2009 distributed by 130 distinct corporate groups.[5]

Figure 2 provides some evidence on the dynamics of unit sales and prices. In each case, we represent the median value during week t. Regarding sales (in thousands of units) we see that a large fraction takes place in the weeks following release. After six months or so, sales are down to a considerably lower level, and they continue declining over time, though at a lower rate. Another noticeable feature of the quantity data is that there are significant "anniversary" effects, namely spikes in quantity sales at around each yearly anniversary from release. In the regressions we present below we include calendar and age fixed effects, which effectively take care of these spikes.

Regarding prices, we notice a decline over time, though at a much lower rate than for unit sales. The median price starts at about $15, and after 1.5 years stabilizes at about $10.

■ **Bundles.** In addition to singles sales, 1,059 bundles (by our estimate) were on the market. We determine an item is a bundle when its name includes the names of different feature films. Typically a bundle consists of two different DVDs; occasionally, three DVDs are included in the same bundle. Bundles are nearly always offered in a mixed-bundling regime, that is, sales of singles titles are also available.[6] Moreover, once a bundle becomes available it is available for the remainder of our sample. This means that t_{xy}, the time when the bundle of x and y is introduced, is a sufficient statistic for the strategy of mixed bundling (of x and y).

5. Our data includes video sales under all formats. Sometimes companies re-release a video title under a different format, e.g., Blu-Ray; we define "new" releases based on the original release date as recorded video, rather than on title-format combinations.
6. There are a few exceptions when a bundle was offered before the second movie title was available as a single.

Table 1
Summary statistics for bundles

Variable	N	mean	sd	p1	p99
Mean age since DVD release	1,059	6.83	5.11	0.14	19.67
Mean user rating	1,058	6.19	1.07	3.05	8.30
Std. dev. of user rating	1,057	0.66	0.56	0.00	2.62
Mean box-office revenue (US$M of 2009)	861	68.23	57.61	0.08	273.89
Std. dev. of box-office revenue	723	34.99	39.64	0.10	191.50
Std. dev. (in 000s days) of release dates	1,059	1.01	1.05	0.00	4.45
Share a distributor (0/1)	1,059	0.98	0.15	0.00	1.00
Share top actors or directors, pooled (0/1)	1,059	0.30	0.46	0.00	1.00
Share top actors (0/1)	1,059	0.26	0.44	0.00	1.00
Share director (0/1)	1,059	0.09	0.29	0.00	1.00
Same genre (0/1)	1,059	0.67	0.47	0.00	1.00
Same language (0/1)	1,059	0.99	0.08	1.00	1.00
Same MPAA rating (0/1)	550	0.68	0.47	0.00	1.00
Same release medium (0/1)	1,059	0.97	0.16	0.00	1.00

Table 1 provides some descriptive statistics for these bundles. Some observations that stand out:

- 98% of all bundles correspond to titles issued by a given studio ("share a distributor").
- 26% of all bundles include movies starring the same lead actor.
- The original release dates of a bundle's component DVDs are typically 3 years apart (1,010 days).[7]

Figure 3 plots the kernel density of bundle release dates, specifically the week of the year when a bundle is released.[8] As can be seen, there are two spikes around Thanksgiving and Christmas, suggesting that one purpose of the bundling strategy is to provide consumers with gift purchasing opportunities. That said, the figure suggests that concentration around holidays is not particularly high. We will return to this later.

Finally, we notice that the average user rating of the titles included in bundles is 6.19, with a standard deviation of 1.07. Compared to this, the standard deviation of the ratings of the titles included in the bundle, 0.66 on average, seems rather small. We regard this as an important observation. One common perception regarding the practice of bundling movies is that a "hit" is used to push a "dud." The simple summary statistics seem at odds with this view: bundles seem to include movies of relatively similar quality (as judged by users).

Are some studios more likely to bundle than others? Figure 4 plots the number of movies and number of bundles by studio. One would expect the relation to be somewhat convex: a studio with n movies can create up to $n\,(n-1)$ different bundles, a number

7. For bundles comprising two titles only, the standard deviation is simply the difference in release dates.
8. Gaussian kernel with 0.05 bandwidth. Similar shapes are obtained for different bandwidth values.

Figure 3
Time of bundle release

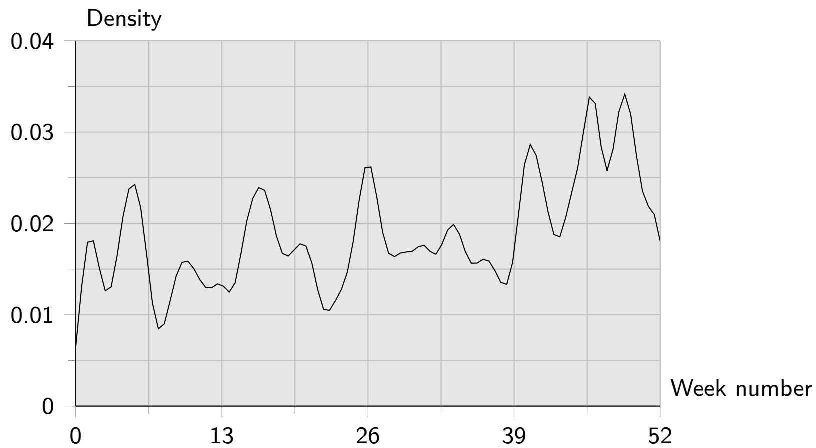


Figure 4
Propensity to bundle by distributor

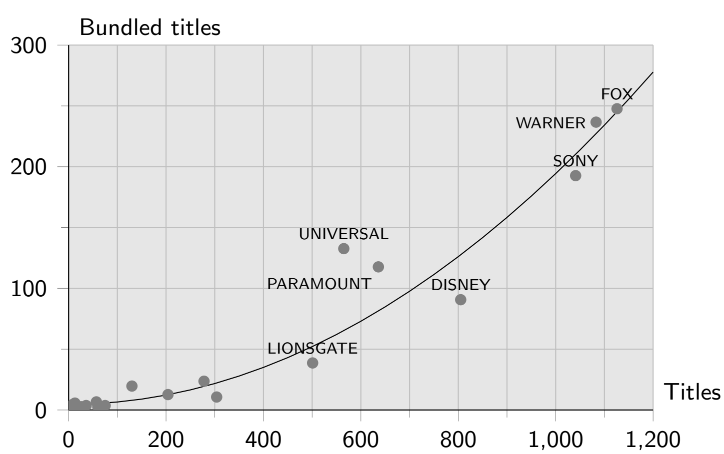


Table 2
Hypothetical and actual bundles

Variable	**Actual (mean)**	**Hypothetical (mean)**	**t-stat. of difference**
Mean user rating	6.19	6.24	1.01
Std. dev. of user rating	0.66	0.94	9.14
Mean box-office revenue (US$M of 2009)	68.23	47.98	-7.71
Std. dev. of box-office revenue	34.99	45.67	4.37
Std. dev. (in 000s days) of release dates	1.01	1.78	14.29
Share a distributor (0/1)	0.98	0.13	-72.12
Share top actors or directors, pooled (0/1)	0.30	0.01	-17.71
Number of actors or directors shared	0.62	0.01	-7.88
Share top actors (0/1)	0.26	0.01	-16.30
Share director (0/1)	0.09	0.00	-8.90
Same genre (0/1)	0.67	0.21	-22.06
Same language (0/1)	0.99	0.90	-9.46
Same MPAA rating (0/1)	0.68	0.42	-7.46
Same release medium (0/1)	0.97	0.99	2.77

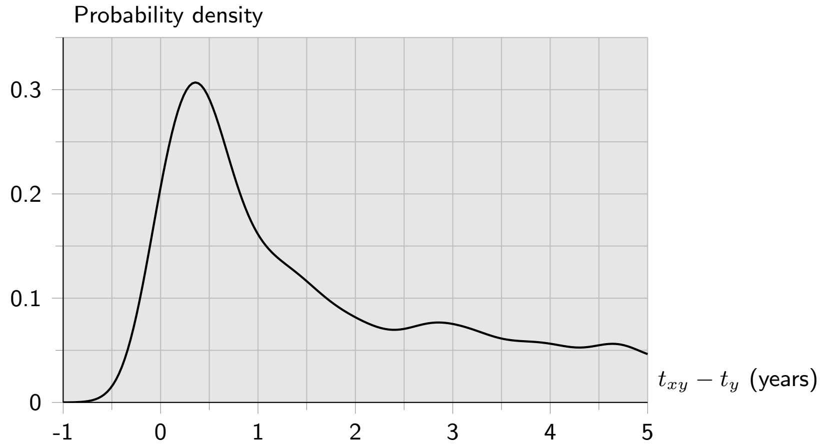
that increases in the order of n^2. In fact, a quadratic curve provides a very good fit for the relation between number of titles and number of titles included in a bundle. Although there are some distributor outliers, the difference from the norm is rather small. We thus conclude that distributor-specific bundling effects are small, beyond the effect of distributor size on the probability of bundling.

In sum, a very preliminary look at the data suggests that bundles are determined by a studio and include movies that are of similar quality and share certain characteristics, specifically each movie's lead talent. We next take a closer, more systematic approach to understand the nature of the studios' bundling strategy.

■ **What movies get bundled and when.** About one quarter of the bundles issued share a leading actor. Is this a high or a low number? In order to get a better feel for the nature of the distributors' bundling strategy, we propose the following exercise: for each bundle xy, we create a hypothetical bundle combining x (i.e., the individual title released earlier) and a randomly selected not ever bundled y' movie; and then compare the average characteristics of these hypothetical bundles to the average characteristics of actual bundles. (We perform a similar analysis pairing y, the later-released title of the actual bundle, with a random x' film that was released earlier than y but never formed part of an actual bundle, and obtain qualitatively the same results as the ones reported here.)

Table 2 presents the results of this exercise. The first column with numbers shows the average values for the actual bundles. The next column corresponds to the hypothetical bundles mentioned in the previous paragraph. Finally, the third column displays the t statistic for the equality of means test. The message is clear: bundles are not random

Figure 5
Kernel density of time elapsed from second release (t_y) to release of bundle (t_{xy})



pairings. Rather, bundles disproportionately combine DVDs of a similar genre, language, MPAA rating, movies with the same director and/or actors, and DVDs that were released at relatively close dates (3 years as opposed to the average of 5).

Two more notes stand out in Table 2. First, bundles do not seem very different in terms of user rating. Second, bundles do differ in terms of box-office revenue: an average bundle includes movies that grossed \$68 million; the corresponding value for a random bundle is \$48 million.

Finally, Figure 5 plots the kernel density of $t_{xy} - t_y$, the time difference, measured in years, between the release of the bundle and the release of the second title included in the bundle. The density is particularly high around zero — and for a good number of titles $t_{xy} = t_y$. However, the right tail is quite thick.

■ **A closer look at prices and the bundling discount.** Figure 6 shows the kernel density estimate of singles prices before and after bundling takes place. Specifically, we compute average prices for a given movie x across all stores and across all weeks in a one-quarter window around the bundling decision. The figure suggests that there is very little difference between the price distributions before and after bundling takes place, except for some shift in mass across different modes of the price distribution: an increase in mass around \$10 and \$13 and a decrease around \$20.

Figure 7 plots the kernel density of the bundling discount for the bundles in our sample, that is,

$$d \equiv p_x + p_y - p_{xy}.$$

We use the average prices across all stores and across all weeks in a one-quarter window around the bundling decision. As can be seen, the average bundling discount is clearly positive. The mode is at around \$4. Strangely enough, we observe cases when the bundling discount is negative. We note, however, that we are working with data that is aggregated across stores. This could therefore be an artifact of aggregation.[9]

9. Moreover, some of our bundles are "special editions" that include additional features, so that the bundle is more than the sum of the parts.

Figure 6
Pre- and post-bundling (single DVD) prices

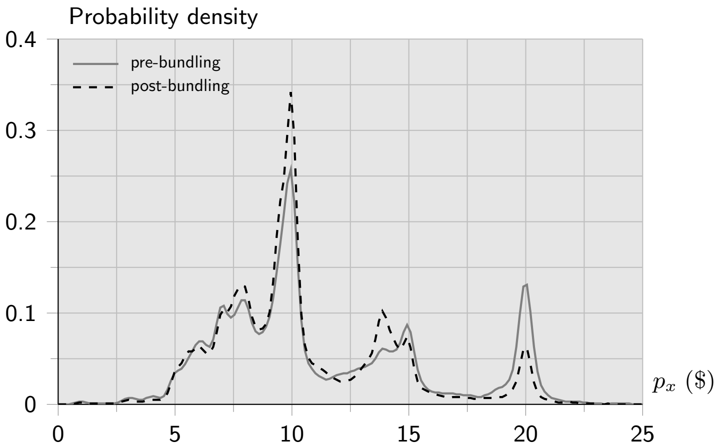


Figure 7
Bundling discount

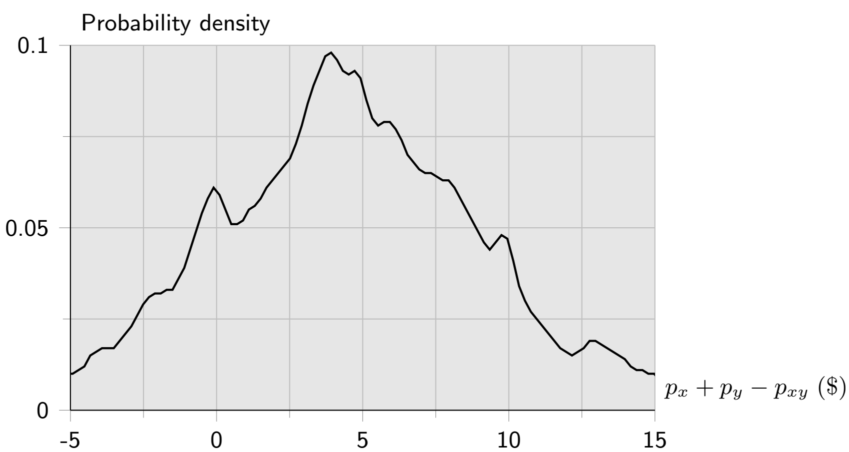
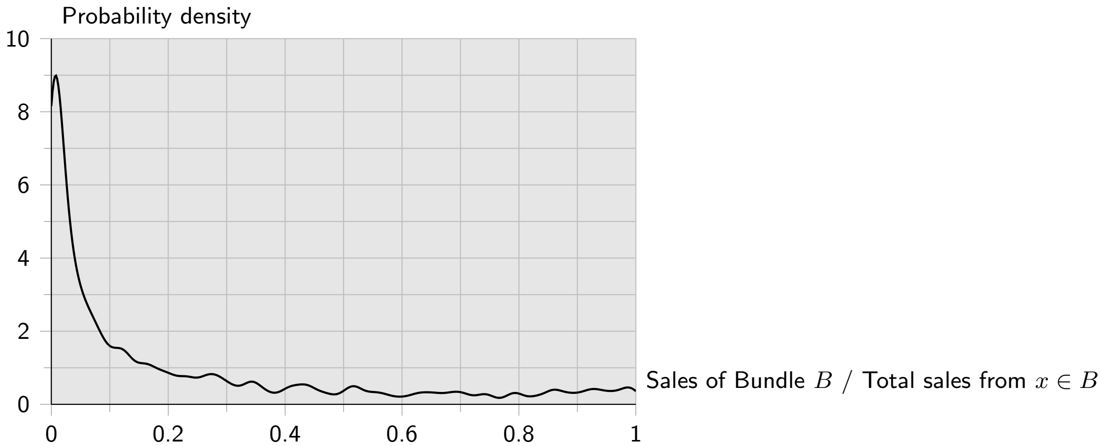


Figure 8
Bundle sales and total sales



The bundles in our sample are instances of mixed bundling (with a handful of exceptions): in addition to the bundle, consumers may purchase the individual titles as well. Naturally, de jure mixed bundling may turn into de facto pure bundling if the bundling discount is so large that no consumer purchases individual titles. One way to measure how close mixed bundling is to pure bundling is to calculate the fraction of total sales of a given title that is obtained through a bundle as opposed to single sales. Figure 8 shows the kernel density of this measure (Gaussian kernel, density bandwidth of .05). As can be seen, there is a substantial fraction of title sales for which bundle sales represent a small fraction of total sales. Aside from this fraction of bundles, the remaining values are distributed approximately uniformly across fraction values all the way to 100%, the case of pure bundling. In other words, while some of our bundles are close to de facto pure bundling (most revenues result from bundle sales) the rule is that of mixed bundling.

To summarize the descriptive evidence so far, we have seen that

- Most sales for single titles take place during the first few weeks.
- Prices drop from about \$15 to about \$10 in 1.5 years.
- There are some "anniversary" effects in sales (though not in prices).
- Most bundles are introduced soon after the second DVD release.
- Bundles originate from the same studio and consist of similar titles (user rating, box-office revenue, lead actor, etc).
- Distributors are equally likely to combine titles into bundles, so that the number of bundles is proportional to the square of the number of available titles.
- Bundling has little effect on the prices of singles.
- The bundling discount is about \$4.

Most of these facts are probably not surprising. Also, we note that bundling is not a device to "push" a bad product with a good one. This runs counter a popular view regarding bundling.

4. Results: gains from bundling

Is bundling a profitable strategy, as predicted by Proposition 1? How much do seller revenues change when bundling is introduced? A naive way of answering this question would be to run a regression of sales revenues on a bundling dummy. However, this would not account for endogeneity. In particular, a typical feature of media products — including DVDs — is that, all else equal, price, quantity and revenues tend to decrease over time. For DVDs, this is shown in Figure 2. In our sample, a bundle is available from time t until the end of the sample period. Given this, a simple regression of revenues on a dummy representing the bundling decision would likely produce a biased estimate, possibly even with the wrong sign.

Our strategy to take these problems into account is to (a) include calendar time and title age fixed effects, and (b) compare revenues with and without bundling around the moment when the bundling decision takes place.

The first step is to assign bundling revenues to individual movie titles. In this way, we are able to continue our analysis at the movie level. Let x and y be two DVD titles and xy the bundle of these two titles. Let b be a dummy variable such that $b = 0$ if no bundle is offered and $b = 1$ if a bundle is offered.[10] We define a series of variables. First, total revenues R^b, before and after bundling takes place.

$$\begin{aligned} R^0 &= p_x^0\, q_x^0 + p_y^0\, q_y^0 \\ R^1 &= p_x^1\, q_x^1 + p_y^1\, q_y^1 + p_{xy}^1\, q_{xy}^1. \end{aligned}$$

Next, we define prorated revenues. These are revenues attributed to a given movie title that forms part of a bundle in addition to the revenues of its sales as a single item. There exist alternative ways to assign these bundle revenues (and as we will report in the robustness section, the alternatives we implemented did not affect our main results). In the case of a two-movie bundle, we propose a simple rule for prorating revenues:

$$\begin{aligned} r_x^0 &= p_x^0\, q_x^0 \\ r_x^1 &= p_x^1\, q_x^1 + \tfrac{1}{2}\, p_{xy}^1\, q_{xy}^1. \end{aligned}$$

Having computed r_x^b in this way, we regress r_x^b on the dummy b as well as a series of other regressors, including in particular calendar and age fixed effects. (Recall that $b = 0$ for all t before a bundle is released and $b = 1$ for all t after a bundle is released.) For each of the bundles in our sample, we include observations at the movie-week level; the sample period is one quarter before and one quarter after the release of the bundle as well as the week of release, totaling twenty seven weeks; this sample is unbalanced, as some titles that form part of a bundle have been on the market for a shorter time. Standard errors in the regressions are clustered at the level of each film.

The results are reported in Table 3. The most important results are shown in the first rows, the ones corresponding to the bundling dummy and its interaction with other variables. First, we notice that, consistent with Proposition 1, the "independent effect" of bundling, estimated in the first model, is .398, that is, an increase in revenues of about

10. Recall that, with rare exceptions, we only observe mixed bundling, that is, when a bundle is offered the single titles are also offered. In a handful of cases, a bundle xy was introduced *before* y was released as a single.

Table 3
Mixed bundling and revenues

Dependent variable	$log(r_x)$	$log(r_x)$	$log(r_x)$	$log(r_x)$	$log(r_x)$	$log(r_x)$	$log(r_x)$
Mixed-bundling regime	0.398***	0.382***	0.331***	0.377***	0.448***	0.452***	0.463***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
...× sequel		0.207**					0.191*
		(0.10)					(0.10)
...× shares top actors or directors			0.204***				
			(0.05)				
...× number of top actors or directors shared				0.024***			0.024***
				(0.01)			(0.01)
...× std.dev. release dates					-0.051**		-0.048**
					(0.02)		(0.02)
...× std.dev. rating of titles						-0.084*	-0.083**
						(0.04)	(0.04)
Stars' box office	0.006	0.006	0.005	0.007	0.007	0.007	0.007
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distributor sales	0.059**	0.059**	0.062**	0.061**	0.059**	0.056**	0.058**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Genre sales	0.088***	0.087***	0.089***	0.088***	0.089***	0.088***	0.088***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Title fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-week dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Title age (in weeks) dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	0.88	0.88	0.88	0.88	0.88	0.88	0.88
N. observations	23462	23462	23462	23462	23462	23462	23462
N. clusters	1128	1128	1128	1128	1128	1128	1128

40%. The dependent variable includes revenues both of the individual title and of the prorated sales of the bundle for that title; alternatively, in an untabulated model that uses as the dependent variable only the revenues of each individual title excluding the revenues from the bundle, the point estimate on the mixed-bundling regime dummy is negative but statistically insignificant. Bundling boosts revenues for the individual title as a whole.

The next models — corresponding to the various columns in Table 3 — consider various possible interaction variables. For example, the second model show that, for bundles that are not sequels, the revenue increase is given by 38%, whereas for sequels such increase is given by $.382 + .207 = 59\%$.

All in all, we consider five different variables that measure the similarity of DVDs included in the same bundle: a dummy for sequels; a dummy for movies that share some top actors or directors; number of top 5 actors plus director shared; standard deviation of release dates; and standard deviation of user rating. Note that the latter two variables (standard deviation of release dates and of user rating) are negative measures of similarity of bundle components.

We have already established that bundles disproportionately include similar titles. The results in Table 3 suggest that, consistent with Proposition 1, the predicted gain from mixed bundling is greater the greater the degree of similarity among the titles included in the bundle. Sharing top talent (at least one actor) is associated with a 20% extra increase in total revenues. Measuring the number of actors in common, we get 2.4% per actor, which together suggests a decreasing marginal effect.

Regarding the standard deviation of release dates, a negative measure of similarity among bundle components, we estimate that a one-standard deviation decrease in the independent variable (greater similarity) is associated with 5.4% higher revenues. Finally, a one-standard deviation decrease in the standard deviation of average user ratings (greater similarity) is associated with 4.7% higher revenues.

As a complement to the results in Table 3, Table 4 shows how the bundling decision is associated with units sold of a single DVD as well as units sold both as a single and as a bundle. The first pair of models suggests that bundling is associated with an increase in total unit sales but with no significant change in singles sales. Moreover, these patterns seem not to vary across sequels and non-sequel bundles. However, as we saw earlier, there is considerable heterogeneity across bundles regarding the importance of bundle sales in total sales. With that in mind, we split the sample of bundles into those where bundles represent an above-median share of total unit sales. The second set of regressions—displayed in the last columns of Table 4—suggests that, for bundles that were relevant for total unit sales, a bundle is associated with an increase in total unit sales (almost a doubling of total unit sales, an increase of $0.427 + .548 = 97.5\%$), whereas the sales of singles drop by about 9.2%.

Our theoretical model, stylized as it is, predicts that bundling leads to lower sales of x and has no effect on the sales of y or on singles pricing. When we distinguish between q_x and q_y (that is, the old and the new DVD) in the regressions of Table 4, we estimate (in unreported regressions) that the effect on q_x is more negative than the effect on q_y. However, the difference in coefficients is not statistically significant. As to prices, Figure 6 is roughly consistent with the prediction of no effects on prices: the pre- and post-bundling densities looks similar, though a closer look suggests that post-bundling singles prices are lower (due to a transfer of density from about \$20 to about \$10). However, we believe this is one aspect where the very stylized nature of the theoretical model (two types) makes

Table 4
Bundling and unit sales ($s_x = 1/n$, where n is the number of titles in the bundle; ϕ is the fraction of total revenues accounted by bundle sales)

	$log(q_x)$	$log(q_x + s_x\, q_{xy})$	$log(q_x)$	$log(q_x + s_x\, q_{xy})$
Mixed-bundling regime	-0.018	0.751***	0.088**	0.427***
	(0.03)	(0.04)	(0.04)	(0.04)
...× sequel	0.030	-0.003	0.047	-0.056
	(0.11)	(0.11)	(0.11)	(0.12)
...× above median ϕ			-0.180***	0.548***
			(0.04)	(0.05)
Stars' box office	0.003	0.008	0.003	0.008
	(0.00)	(0.01)	(0.00)	(0.01)
Distributor sales	0.016	0.047	0.015	0.050
	(0.04)	(0.04)	(0.04)	(0.04)
Genre sales	0.087***	0.092***	0.087***	0.091***
	(0.02)	(0.03)	(0.02)	(0.03)
Title fixed effects	Yes	Yes	Yes	Yes
Year-week dummies	Yes	Yes	Yes	Yes
Title age (in weeks) dummies	Yes	Yes	Yes	Yes
Adjusted R2	0.88	0.82	0.88	0.82
N. observations	53947	53947	53947	53947
N. clusters	1489	1489	1489	1489

this sort of prediction problematic. For example, the prediction of no effect on p_x clearly depends on there being only two consumer types. More generally, one might expect singles prices to decrease with the introduction of a bundle (as Figure 6 suggests).

■ **Robustness checks.** Our analysis of the relation between bundling and total revenues and unit sales as well as the relevance of correlation of movie characteristics shows a number of economically and statistically significant results. We performed a series of robustness checks on these results. First, our analysis has been conducted at the movie title level, something we do by prorating bundle revenues and unit sales to the constituting bundle component titles. In the process, we treat symmetrically all titles of the bundles. One might ask whether the first title in the bundle performs differently in a systematic manner. We split our sample into x movies (first release) and y (subsequent releases), finding no significant differences in the various regression coefficients.

A second important assumption in this process is to assign a $s_x = \frac{1}{n}$ share to each of the titles in a bundle (with $n = 2$ for almost all bundles). An alternative is to prorate bundles sales according to pre-bundling sales levels, where the pre-bundle sales share is calculated immediately before the bundle release date, thus containing information about the relative popularity of individual titles at that moment:

$$r_x^1 = p_x^1\, q_x^1 + s_x^0\, p_{xy}^1\, q_{xy}^1$$

where

$$s_x^0 \equiv \frac{p_x^0\, q_x^0}{p_x^0\, q_x^0 + p_y^0\, q_y^0}.$$

The results using the pre-bundling sales levels for the prorate are similar to the ones we reported using $s_x = \frac{1}{n}$. This is not entirely surprising: as we saw in Section 3, bundled movies tend to be similar in various characteristics, including user reviews and box-office performance. To further account for differences in quality characteristics, we also prorated bundle sales using the pre-bundling *relative prices* of the individual items immediately before the release date of the bundle, or their box-office revenues on the theatrical market, or their user ratings. In all these alternative models we obtained qualitatively the same results.

We considered a number of variations on the models presented in Table 3 and 4. For example, we included additional interaction variables in the revenue analysis such as a Christmas dummy (insignificant effect) and the average rating of the bundle component titles (again, insignificant effect). We also estimated separately the effects of bundling on unit sales by type of store (e.g., online vs. offline sellers). The results do not change in any considerable way.

Finally, as shown in Figure 3 we do observe some small spikes in the density of bundle release dates around Thanksgiving and Christmas. However, the pattern of bundle release is relatively uniform along the calendar year. That said, we reran the above regressions restricting the analysis to the sample of weeks 1 to 45 in the calendar year, thus excluding bundles released during the holiday season. The coefficient estimates are very close to those in Tables 3 and 4, as is their statistical significance.

5. Discussion and concluding remarks

By bringing together goods whose valuations are negatively correlated, bundling helps "homogenizing" total valuations and thus extract a greater share of consumer surplus. This is, to a great extent, the "conventional wisdom" regarding bundling. However, anecdotal evidence — and systematic evidence from examples such as DVD sales — suggests that bundles include goods whose consumer valuations are positively correlated, which seems to contradict the conventional wisdom.

Following Derdenger and Kumar (2013), we argue that, in an intertemporal price discrimination context, even if consumer preferences across bundle components are positively correlated, bundling can be a revenue-increasing strategy. In fact, bundling works *precisely* because consumer valuations are positively correlated. We develop a simple theoretical model consistent with this narrative of bundling with durable goods. The model implies two testable predictions. First, bundling an "old" durable with a recently-released good increases seller revenues. Second, such increase is greater the more similar the two products are. We test our predictions on data from DVD sales in the 2000s. The results provide strong support for the theoretical propositions.

There are, admittedly, other explanations for profitable bundling with positively correlated valuations. For instance, as mentioned in Section 1, Gandal et al. (2018) show that, under *pure* bundling, bundling might have a market-expansion effect because it "fattens" the tail of the distribution of valuations (variance-increasing effect). However, these alternative explanations do not address the dynamic effects we consider in our theoretical model.

We believe the phenomenon we characterize in this paper has relevance beyond the examples suggested in Sections 1 and 3. Many media products share several properties with the DVD industry. For example, a reader of our paper reports that, when looking for a particular Arthur Miller play, he received an offer to the effect that for an extra \$3, he could buy a collection of Miller plays which included the play he was looking for. Similarly, music compilations by artist or by genre can also be interpreted as a form of mixed bundling targeted at low-valuation buyers. More important, bundling hardware and software is a practice that shares the features of our model.

References

Adams, W. and Yellen, J.L. "Commodity Bundling and the Burden of Monopoly." *Quarterly Journal of Economics*, Vol. 90 (1976), pp. 475–498.

Bulow, J.I. "Durable-Goods Monopolists." *Journal of Political Economy*, Vol. 90 (1982), pp. 314–332.

Cabral, L. and Natividad, G. "Cross-Selling in the US Home Video Industry." *Rand Journal of Economics*, Vol. 47 (2016), pp. 29–47.

Chen, Y. and Riordan, M.H. "Profitability of Product Bundling." *International Economic Review*, Vol. 54 (2013), pp. 35–57.

Chu, C.S., Leslie, P., and Sorensen, A. "Bundle-Size Pricing as an Approximation to Mixed Bundling." *American Economic Review*, Vol. 101 (2011), pp. 263–303.

Coase, R.H. "Durability and Monopoly." *The Journal of Law & Economics*, Vol. 15 (1972), pp. 143–149.

Crawford, G.S. and Yurukoglu, A. "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, Vol. 102 (2012), pp. 643–85.

Derdenger, T. and Kumar, V. "The Dynamic Effects of Bundling as a Product Strategy." *Marketing Science*, Vol. 32 (2013), pp. 827–859.

Elberse, A. and Oberholzer-Gee, F. "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales." *Marketing Science Institute*, Vol. 4 (2007), pp. 49–72.

Gandal, N., Markovich, S., and Riordan, M.H. "Ain't it 'Suite'? Bundling in the PC Office Software Market." *Strategic Management Journal*, Vol. 39 (2018), pp. 2120–2151.

Gentzkow, M. "Valuing New Goods in a Model with Complementarity: Online Newspapers." *American Economic Review*, Vol. 97 (2007), pp. 713–744.

Ho, K., Ho, J., and Mortimer, J.H. "The Use of Full-line Forcing Contracts in the Video Rental Industry." *American Economic Review*, Vol. 102 (2012), pp. 686–719.

Johnson, J.P. and Myatt, D.P. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review*, Vol. 96 (2006), pp. 756–784.

McAfee, R.P., McMillan, J., and Whinston, M.D. "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values." *The Quarterly Journal of Economics*, Vol. 104 (1989), pp. 371–383.

Mortimer, J.H. "Vertical Contracts in the Video Rental Industry." *The Review of Economic Studies*, Vol. 75 (2008), pp. 165–199.

Stigler, G.J. "United States v. Loew's, Inc.: A Note on Block Booking." *Supreme Court Review*, Vol. 152 (1963).

Stokey, N. "Rational Expectations and Durable Goods Pricing." *Bell Journal of Economics*, Vol. 12 (1981), pp. 112–128.