Mixed Bundling in Retail DVD Sales: Facts and Theories

Luís Cabral
New York University

Gabriel Natividad
Universidad de Piura

July 2018

Abstract. Many DVD titles are sold in retail stores in bundles, typically a bundle of two different titles with common characteristics: same lead actor/actress, same director, same genre, etc. This suggests that consumer valuations are positively correlated across the bundle components, which in turn runs counter to the received wisdom that bundling is most profitable when valuations are negatively correlated.

In this paper, we propose a solution to this puzzle, one that is based on the observation that DVDs are sequentially released durable goods. At the time the second title is released, it is likely that high-valuation buyers will have bought the first one. For this reason, even though *ex-ante* valuations are positively correlated, *ex-post* — that is, at the time the second title is released — valuations are negatively correlated.

We provide sufficient conditions such that mixed bundling increases revenues and the revenue increase is greater the more positively correlated valuations are. We also provide empirical confirmation of this prediction as well as an independent estimate from a calibrated analytical model.

Cabral: Stern School of Business, New York University; luis.cabral@nyu.edu. Natividad: Universidad de Piura; gabriel.natividad@udep.pe. We are grateful to 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

Many retail stores, such as Walmart or Kmart, sell DVDs of previously released movies. In some cases, DVD titles are sold in bundles, typically a bundle of two different titles. In addition to the bundle, buyers can also choose to purchase the individual titles separately (in other words, it is a case of mixed bundling).

At least since Stigler (1963), the practice of bundling movies has been considered a form of second-degree price discrimination that takes advantage of the negative correlation in buyer valuations. In Stigler's words, "the simplest plausible explanation [for the practice of bundling] is that some buyers would prize one film much more relative to the other" (p. 153). Stigler's (1963) seminal contribution was continued by many authors who have studied the conditions under which bundling is a revenue-increasing strategy. 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.

Crawford and Yurukoglu's (2012) evidence from the U.S. cable industry seems largely consistent with this view. By contrast, the practice of DVD bundling seems largely *inconsistent* with it. Typically, bundled DVD titles have one or more elements in common: the same lead actor/actress, the same director, the same genre, etc. (They are also owned by the same distributor). For example, Universal Pictures' *The Scorpion King*, starring Dwayne Johnson, was released in 2002. In 2003, Universal released another DVD, *The Rundown*, starring the same lead actor. Soon after, retail stores started selling a bundle comprising *The Scorpion King* and *The Rundown*. To the extent that similarity of characteristics is associated with correlation of valuations, this presents a puzzle: if negative correlation of valuations (or "sufficiently limited positive dependence") is the basis of a successful bundling strategy, then why do distributors chose bundles the way they do?

In this paper, we propose a solution to this puzzle. DVDs — just as many other media products — have several distinct characteristics: they are durable goods, they are released sequentially, and there is a great number of different titles available. Two DVDs that share several characteristics are likely to be similarly valued by viewers. However, at the time the second title is released, it is likely that high-valuation buyers will have bought the first one. For this reason, even though *ex-ante* valuations are positively correlated, *ex-post* — that is, at the time the second title is released — valuations are negatively correlated: buyers who have a high valuation for the second title are likely to have a low valuation for the first one — because they have already purchased it before.

In Section 2 of this paper we present a simple two-period, two-type model that formalizes this intuition. We provide sufficient conditions such that the gain from mixed bundling is positive and increasing in the degree of similarity of the goods bundled (and highest when the ex-ante correlation of valuations is perfect).

In Section 3 we analyze a dataset comprising a substantial portion of DVD sales (prices and quantities) in the US from 2000–2009. (Our dataset includes all titles sold as a DVD but not all of the selling stores.) This analysis serves several purposes. First, we document the extent to which bundles are biased towards selecting similar titles (they are). Second, by means of a simple differences 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, in accordance with our theory, we estimate that the gain from mixed bundling is greater the greater the similarity between bundled titles.

In order to assess the value of our theory in explaining the data, in Section 4 we consider an extension of our simple model to the continuous-type case. The model is calibrated with the median values from our sample. The calibrated model confirms the basic result from the two-type model: the gains from mixed bundling are positive and increasing in the degree of correlation in valuations. However, the estimated values of gains are lower than in the reduced-form regressions: from 17% if valuations are independent to 28% if they are perfectly correlated. We also estimate that the optimal bundling discount — keeping the prices for singles at the sample mode — is about $10, substantially more than the median bundling discount in our sample, about $5. Despite the differences in values between the reduced-form model and the calibrated analytical model, overall we find the evidence corroborative of the idea of mixed bundling of sequentially released durable goods as a form of second-degree price discrimination. We also discuss reasons for the divergence in estimates between the two approaches.

■ **Related literature.** We are not aware of many economics papers that estimate the effects of bundling empirically. Gandal et al. (2015) directly address the issue of correlation of preferences. They estimate a discrete-choice model of software demand and apply it to the PC office software market in the 1990s. By simulating various hypothetical market structures, they find that greater correlation in preferences enhances the profitability of bundling due to the interaction of a market expansion effect and a suite bonus effect.

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). On the whole, mean consumer and total surplus change by an estimated -5.4 to 0.2 percent and -1.7 to 6.0 percent, respectively.

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.

Arguably, the empirical 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. However, the key to their result is not sequential releases (as in our case) but rather product differentiation in software.

2. Theory

Our paper is motivated by the apparent puzzle that distributors bundle movies with similar characteristics, which presumably implies that consumer valuations are highly correlated. In this section, we propose a two period, two-type model based on this observation. The model fits particularly well two features of DVDs (and many other markets): durability and sequential release.

This section is structured into two parts. First, we consider the case when consumer valuations are perfectly correlated across products. This limiting case is interesting because, in a static context, bundling has no effect on revenues when valuations are perfectly correlated. By contrast, we show that, in a sequential-release context, bundling has a positive effect on revenues. Second, we show that the gains from bundling are increasing in the degree of similarity across products.

■ **Basic model and intuition.** Consider a seller with two goods that are 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. Throughout this subsection we assume that

$$\underline{u}/\overline{u} > \max\left\{\tfrac{1}{2}, \alpha\right\} \tag{1}$$

Typically, the economic analysis of bundling assumes a static framework where the seller offers a set of products at a given moment of time either as single products or in the form of a bundle. As mentioned in the previous section, media products such as movie DVDs have the important characteristic of being released sequentially over time. This adds an important element to the economic analysis of bundling: when two sequentially introduced products are bundled together — a recently released one and a not-so-recently released one — some buyers may already have purchased the earlier-released product, which in turn affects the relative demand for the new product and the bundle that includes the old product.

Our model of sequentially released products considers two products, x and y; and two time periods, $t = 1$ and $t = 2$. Product x is released at $t = 1$ and Product y at $t = 2$. This means that Product x can be purchased at $t = 1$ or $t = 2$, whereas product y can only be purchased at time $t = 2$. Let p_{xt} be product x's price at time t and p_{y2} 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}.

We assume that consumers are myopic, specifically, they do not consider future price changes or bundling offers in their current purchase decisions. Considering the relatively small changes in price, as well as the unpredictability of future releases, we believe this assumption fits consumer behavior in DVD and related markets.

We first consider the case when buyer valuations are perfectly correlated: a fraction α of buyers has high valuation for both products, whereas a fraction $1 - \alpha$ has low valuation for both products. Notice that, if the seller's problem is atemporal — that is, both products are offered at the same time — then bundling has no effect on seller revenues: the seller's optimal strategy is either to set both prices at $\underline{u}$ or both prices at $\overline{u}$, depending on whether $\underline{u}$ is greater or smaller than $\alpha\,\overline{u}$. By contrast, under sequential product release, bundling strictly increases the seller's payoff:

Proposition 1. *Suppose buyer valuations are perfectly correlated. In equilibrium, the seller is strictly better off by offering a bundle at $t = 2$.*

The complete proof of this and the remaining results is included in the Appendix. We show that, at $t = 1$, high-valuation buyers purchase product x; and at $t = 2$ high-valuation buyers

Figure 1
Imperfectly correlated valuations

	H	L
H	$\alpha\big(\alpha(1-\rho)+\rho\big)$	$\alpha(1-\alpha)(1-\rho)$
L	$\alpha(1-\alpha)(1-\rho)$	$(1-\alpha)\big((1-\alpha)(1-\rho)+\rho\big)$

purchase product y, whereas low-valuation buyers purchase the xy bundle. Specifically, the seller offers a bundle xy for a price $p_{xy} = \underline{u} + \underline{u}$ to attract low-valuation buyers, who have not purchased product x at $t = 1$; and sets $p_{y2} = \overline{u}$ to attract high-valuation buyers. Since $\overline{u} < 2\,\underline{u}$, high-valuation buyers prefer to purchase product y rather than the bundle xy. In other words, sequentiality of sales eases up the high-valuation buyers' incentive-compatibility constraint.

■ 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.

Figure 1 depicts one possible parameterization of joint valuations. The parameter ρ functions as an indicator of correlation. The perfect-correlation case we considered before corresponds to $\rho = 1$, whereas $\rho = 0$ implies independent valuations. Proposition 1 refers to the case when $\rho = 1$. The next result corresponds to the case when ρ is in the neighborhood of 1.

Proposition 2. *In the neighborhood of $\rho = 1$, the seller's gain from bundling is strictly increasing in ρ.*

The idea is that, if ρ is in neighborhood of $\rho = 1$, then the optimal solution remains the same. This is so because the inequalities in Proposition 1 are strict. The effect of lowering ρ away from 1 is therefore an effect on payoffs, not on the nature of firm strategy. Specifically, when $\rho < 1$ we have two new types of buyers. First some HH types become HL types; this shift implies a loss of $t = 2$ seller revenue. Second, some LL types become LH types. Their purchase pattern is the same and leads to the same seller revenue, although their buyer surplus is greater. Finally, this also implies that the revenue loss is increasing in $1 - \rho$.

■ Alternative timing assumptions. As often happens with theoretical models, ours makes several simplifying assumptions which we expect help capture the essential features of the data. One such assumption is that bundles are released at the time when the second product is released. The idea is that, by the time y and xy are released, buyers already have had the option of purchasing x. Thus, even if valuations were positively correlated to begin with, at the time of choice between x, y and xy, they are not.

Our opening example of DVD releases (*The Scorpion King* and *The Rundown*) seems to fit this pattern: the bundle was introduced very soon after the second release. However, Warner Bros.' *The Pelican Brief* and *Conspiracy Theory*, both starring Julia Roberts, started selling as a bundle in 2003, but both titles were released as singles much earlier.

Strictly speaking, examples like *The Pelican Brief/Conspiracy Theory* run against our model's timing assumption. However, we believe the model still captures — as a reduced form — the critical feature we attempt to characterize: There are multiple titles in the market and consumers are not aware of all of them at all times. By the time the *The Pelican Brief/Conspiracy Theory* bundle was released, several buyers were exposed to the possibility of purchasing one of the singles. Our model predicts that high-valuation buyers would then make a purchase, whereas low-valuation buyers would not; so that, when the bundle is released, valuations of xy and y are negatively correlated, as desired, even though (ex-ante) valuations of x and y are positively correlated.

3. Empirical evidence and analysis

The setting for our empirical study is the U.S. home video sales industry during the period 2000–2009.[1] In essence, the video 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.[2] While distributors are large and in small number, retailers range from fairly small specialty stores to larger retail outlets such as Amazon.com.[3]

■ **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.[4]

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.

1. 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.
2. Cabral and Natividad (2016) focuses on the wholesale segment of the industry, whereas this paper focuses on the retail segment.
3. Upstream, distributors obtain content from a series of industries such as feature film, TV and cable producers.
4. 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.

Figure 2
Median sales (units) and price over time

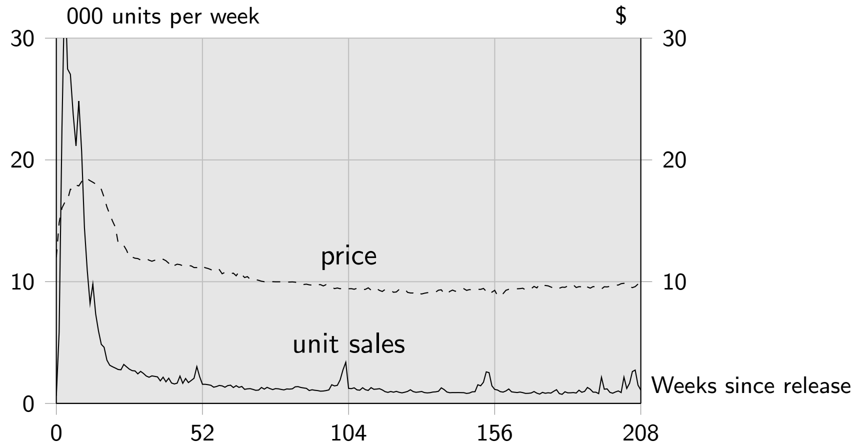
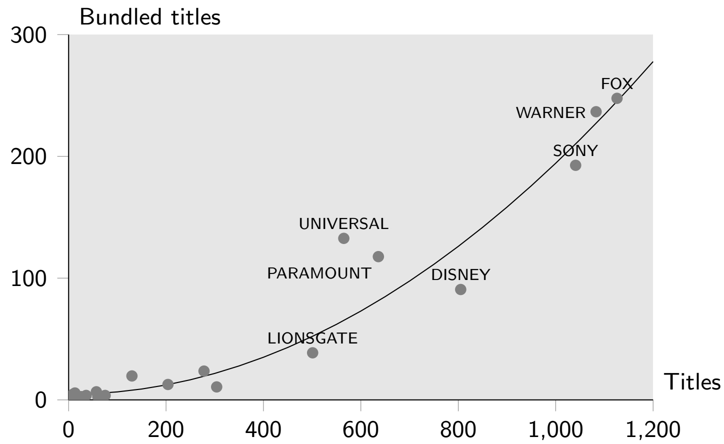


Table 1
Summary statistics for bundles

variable	N	mean	sd	p1	p99
Mean age since DVD release	1,059.00	6.82	5.10	0.14	19.67
Mean user rating	1,058.00	6.19	1.07	3.05	8.30
Std dev user rating	1,057.00	0.66	0.56	0.00	2.62
Mean box office revenue (US$M of 2009)	861.00	68.23	57.61	0.08	273.89
Std dev box office revenue	723.00	34.99	39.64	0.10	191.50
Std dev (in 000s days) of release dates	1,059.00	1.01	1.05	0.00	4.45
Share a distributor (0/1)	1,059.00	0.98	0.15	0.00	1.00
Share top actors or directors, pooled (0/1)	1,059.00	0.31	0.46	0.00	1.00
Share top actors (0/1)	1,059.00	0.26	0.44	0.00	1.00
Share director (0/1)	1,059.00	0.09	0.29	0.00	1.00
Same genre (0/1)	1,059.00	0.67	0.47	0.00	1.00
Same language (0/1)	1,059.00	0.99	0.08	1.00	1.00
Same MPAA rating (0/1)	550.00	0.68	0.47	0.00	1.00
Same release medium (0/1)	1,059.00	0.97	0.16	0.00	1.00

Figure 3
Propensity to bundle by distributor



■ **Bundles.** In addition to singles sales, 1,059 bundles (by our estimate) were placed on sale.[5] 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).

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]

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 3 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 that increases in the order of n^2. In fact, a quadratic curve provides a very good fit for

5. We determine an item is a bundle when its name includes the names of different feature films.
6. There are a few exceptions when a bundle was offered before the second movie title was available as a single.
7. For bundles comprising two titles only, the standard deviation is simply the difference in release dates.

Table 2
Hypothetical and actual bundles

Variable	Actual	Hypo	Diff t
Mean user rating	6.19	6.24	1.12
Std dev user rating	.66	.9	8.39
Mean box office revenue (US of 2009)	68.23	46.59	-8.55
Std dev box office revenue	34.99	46.16	4.69
Std dev (in 000s days) of relase dates	1.01	1.7	13.11
Share a distributor (0/1)	.98	.13	-73.49
Share an actor or director (0/1)	.31	.01	-18.01
Number of actors or directors shared	.63	.01	-8.04
Share an actor (0/1)	.26	.01	-16.18
Share a director (0/1)	.09	0	-8.9
Same genre (0/1)	.67	.22	-21.91
Same language (0/1)	.99	.9	-9.87
Same MPAA rating (0/1)	.68	.38	-8.95
Same release medium (0/1)	.97	.99	2.77

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.

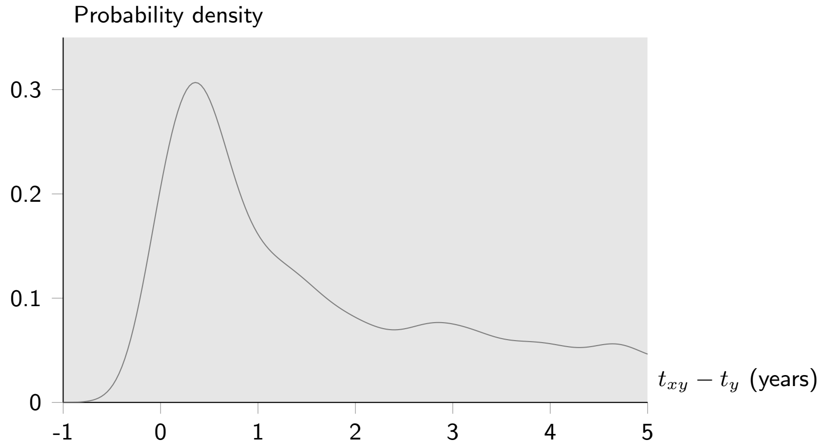
■ **What movies get bundled and when.** About one quarter of the bundles issued share a leading actor. Is this a high or a low number? In order to get a better feel for the nature of the distributors' bundling strategy, we propose the following exercise: for each bundle xy, we create a hypothetical bundle combining x and a randomly selected not bundled y' movie; and then compare the average characteristics of these hypothetical bundles to the average characteristics of actual bundles.

Table 2 presents the results of this exercise. The first column with numbers shows the average values for the actual bundles. The second column corresponds to the hypothetical bundles mentioned in the previous paragraph. Finally, the third column displays the t statistic for the equality test.

The message is clear: bundles are not random 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 that stand out of Table 2. First, bundles do not seem very different in

Figure 4
Kernel density of time elapsed from second release (t_y) to release of bundle (t_{xy})



terms of user rating. They do differ, however, in terms of box-office revenue: an average bundle includes movies that grossed \$68 million; the corresponding value for a random bundle is \$26 million.

Finally, Figure 4 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 5 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 6 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 artifice of aggregation.[8]

All of 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

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

Figure 5
Pre- and post-bundling (single DVD) prices

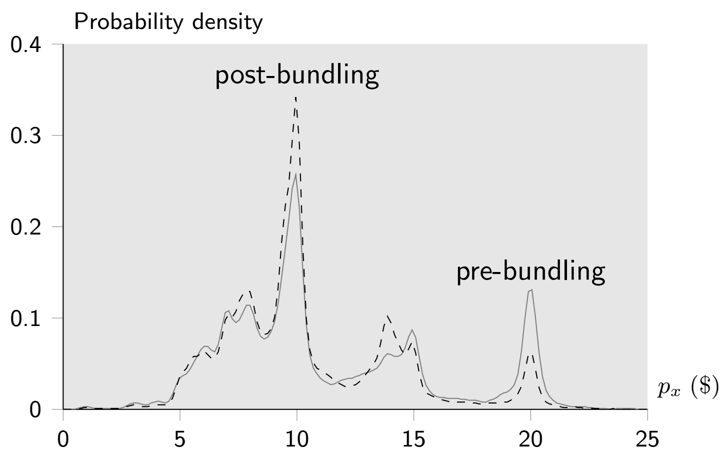


Figure 6
Bundling discount

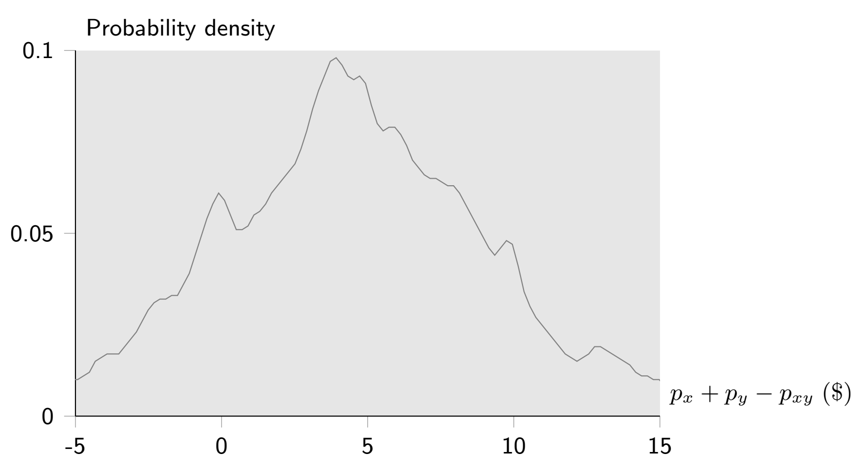
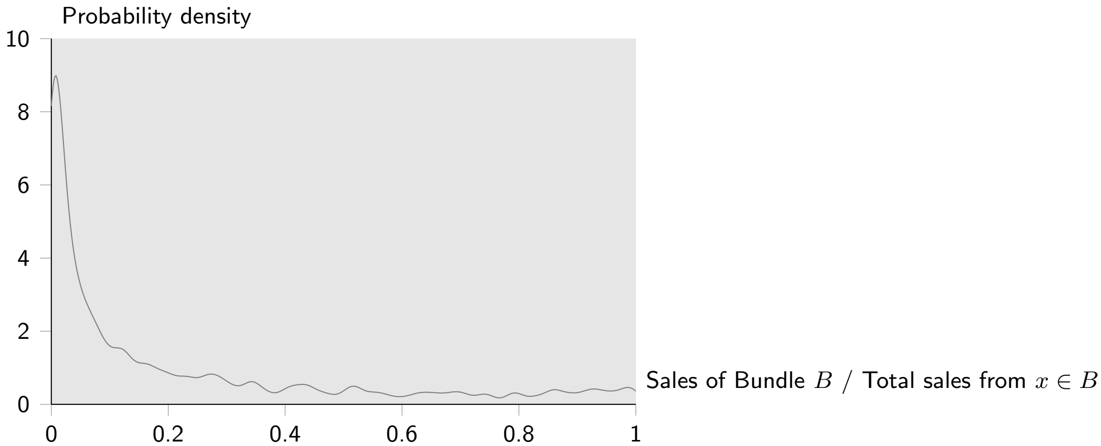


Figure 7
Bundle sales and total sales



close mixed bundling is to pure bundling is to measure the fraction of total sales of a given title that are obtained through a bundle as opposed to single sales. Figure 7 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 (except of course the fact that sellers bundle similar products, though Section 2 provides an answer to this potential puzzle). 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. For example, a compilation of "12 Ways To Sell What's Not Selling" includes "Try bundling the slower-moving product with a better seller."[9]

9. See http://www.profitguide.com/industry-focus/retail/turn-static-stock-into-sales-42620, visited on October 20, 2016.

■ **Measuring the gains from bundling.** Is bundling a profitable strategy? How much do seller revenues change when bundling is introduced? A naive way of answering this question is 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 and 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, including those from bundle sales:

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

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.

The results can be seen 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 the "independent effect" of bundling, estimated in the first model, is .395, that is, an increase in revenues of about 40%. The next few models 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 .380+.2=58%.

All in all, we consider five different variables that measure the similarity of DVDs included in the same bundle: sequels; movies 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 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

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	0.395***	0.380***	0.325***	0.373***	0.443***	0.447***	0.458***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
...× sequel		0.200*					0.186*
		(0.10)					(0.10)
...× shares top actors or directors			0.201***				
			(0.05)				
...× number of top actors or directors shared				0.024***			0.024***
				(0.01)			(0.01)
...× std.dev. release dates					-0.050**		-0.048**
					(0.02)		(0.02)
...× std.dev. rating of titles						-0.082*	-0.081*
						(0.04)	(0.04)
Stars' box office	0.007*	0.007*	0.006	0.008*	0.008*	0.007*	0.008*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distributor sales	0.060**	0.060**	0.062**	0.062**	0.060**	0.058**	0.059**
	(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. obs	23391	23391	23391	23391	23391	23391	23391
N. clusters	1126	1126	1126	1126	1126	1126	1126

Table 4
Bundling and unit sales

	$log(q_x)$	$log(q_x + s_x q_{xy})$	$log(q_x)$	$log(q_x + s_x q_{xy})$
Mixed-bundling regime	-0.013	0.755***	0.096**	0.431***
	(0.03)	(0.04)	(0.04)	(0.04)
...× sequel	0.058	0.011	0.067	-0.016
	(0.11)	(0.12)	(0.11)	(0.12)
...× above median ϕ			-0.182***	0.544***
			(0.04)	(0.05)
Stars' box office	0.006	0.011**	0.006	0.011**
	(0.00)	(0.01)	(0.00)	(0.01)
Distributor sales	0.008	0.035	0.007	0.037
	(0.04)	(0.04)	(0.04)	(0.04)
Genre sales	0.082***	0.087***	0.083***	0.085***
	(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. obs	53925	53925	53925	53925
N. clusters	1489	1489	1489	1489

one actor) is associated with a 20% extra increase in total revenues. Measuring the number of common actors, we get 2.4% per common 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 5.3% higher revenues. Finally, a one-standard deviation decrease in the standard deviation of average user ratings (greater similarity) is associated with 4.9% 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 bundles are 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.

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-mean share of total unit sales. The second set of regressions 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 43.1+54.4=97.5%), whereas the sales of singles drops by about 8.6%.

■ **The right bundle.** Our results allow us to do a simple experiment to test the extent to which sellers follow a good bundling strategy. A lower bound on such a test is to compare actual choices to random choices. Specifically, we create a series of hypothetical bundles that randomly match one of the DVD components of the bundle with a DVD that was

Table 5
Hypothetical and actual bundles, revised

variable	Actual	Hypo	Diff t
Mean user rating	6.19	6.22	.62
Std dev user rating	.66	.9	8.08
Mean box office revenue (US of 2009)	68.23	65.52	-1
Std dev box office revenue	34.99	30.29	-2.35
Std dev (in 000s days) of relase dates	1.01	1.72	13.58
Share a distributor (0/1)	.98	.18	-61.36
Share a top actor or director (0/1)	.31	.01	-17.53
Number of top actors or directors shared	.63	.01	-7.88
Share a top actor (0/1)	.26	.01	-15.95
Share a director (0/1)	.09	0	-8.26
Same genre (0/1)	.67	.21	-22.23
Same language (0/1)	.99	.98	-3.04
Same MPAA rating (0/1)	.68	.33	-11.01
Same release medium (0/1)	.97	.99	2.69

not bundled, and then compare the predicted revenue increase from such bundles to the predicted revenue increase from actual bundles.

Earlier we constructed a set of such hypothetical bundles and noticed significant differences, including differences in box-office revenues between bundle titles and average titles. In order to correct for this possible source of bias, we rebuild our hypothetical bundle set by matching titles with similar box-office revenue. Specifically, we create a matched hypothetical set of bundles as follows. For each actual bundle, we replace the second DVD (title y in our previous notation) with a title not bundled that matches y in box-office revenue but is otherwise randomly chosen.

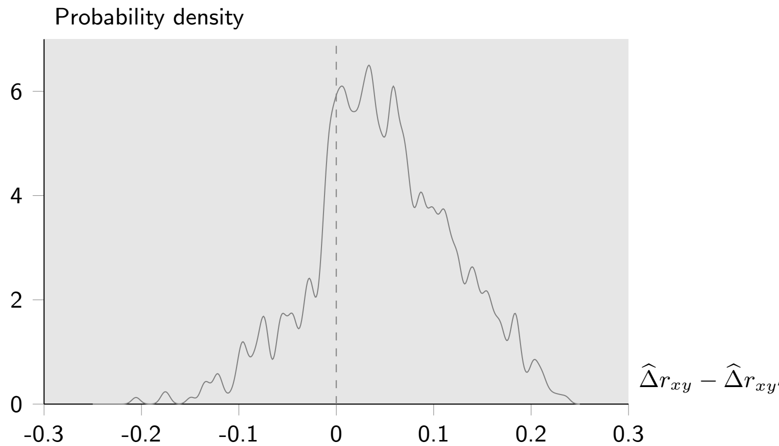
The summary statistics from such revised set of hypothetical bundles is shown in Table 5. We create this revised sample by matching on the variable "Mean box office revenue (US\$M of 2009)." Not surprisingly, the difference between hypothetical and actual bundles is insignificant for this variable. However, for the remaining variables we still observe significant differences (high t ratios) as in Table 2.

Equipped with these two sets of bundles, we then construct a variable $\Delta = \widehat{\Delta} r_{xy} - \widehat{\Delta} r_{xy'}$ equal to the predicted revenue difference between the two bundles associated with each movie x that is included in a bundle. Finally, we proceed to plot the kernel density of this variable (we consider a Gaussian kernel with density bandwidth .05).

The results, shown in Figure 8, show that, on average, sellers do better than issuing random bundles. Specifically, the null hypothesis that $\mu = 0$ is rejected with $p < .01$.

■ **Robustness checks.** Our results regarding the relation between bundling and total revenues, 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

Figure 8
Bundles and pseudo-bundles



checks on these results. First, our analysis is done 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). We observe 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 ($n = 2$ for almost all bundles). An alternative is to prorate bundles sales according to pre-bundling sales sales levels:

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

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 result are similar to the ones we obtain with $s_x = \frac{1}{2}$. 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.

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

Finally, we also attempted some matching analysis. Specifically, we considered all the within-distributor combinations of feature-film-based DVD items, totaling 2.2 million hypothetical combinations of possible bundles. We then ran a probit model of which of these combinations were actual bundles, using as observables for this probit model the following variables: average using rating, standard deviation of the user rating, a dummy for whether there were shared actors, a dummy for whether there were common directors, the standard deviation of release dates, and the average box office revenue of the films. This probit model yielded some very close counterfactuals, based on observables, of the actual bundles. In each case, we took only the actual bundle and its single closest counterfactual hypothetical bundle. For these pairs, we computed the performance (in terms of dollar sales and unit sales) of the component DVD items before and after the release of the latest individual DVD item of the bundle, either real or hypothetical. We did not find statistically significant differences between the performance of the real bundles and the performance of the hypothetical bundles.

■ **Summary.** We summarize our empirical analysis as follows:

- Mixed bundling is associated with an increase in revenues
- Bundles typically consists of DVD titles that are similar to each other
- The gains from bundling are greater when the bundle components are more closely related to each other
- Our empirical model is consistent with seller optimal bundling decision: higher revenues from actual bundles than from hypothetical random bundles

In the next section, we calibrate an continuous-type extension of our basic model so as to obtain alternative estimates of the gains from bundling; and so as to perform a number of counterfactual exercises.

4. Calibration and simulation

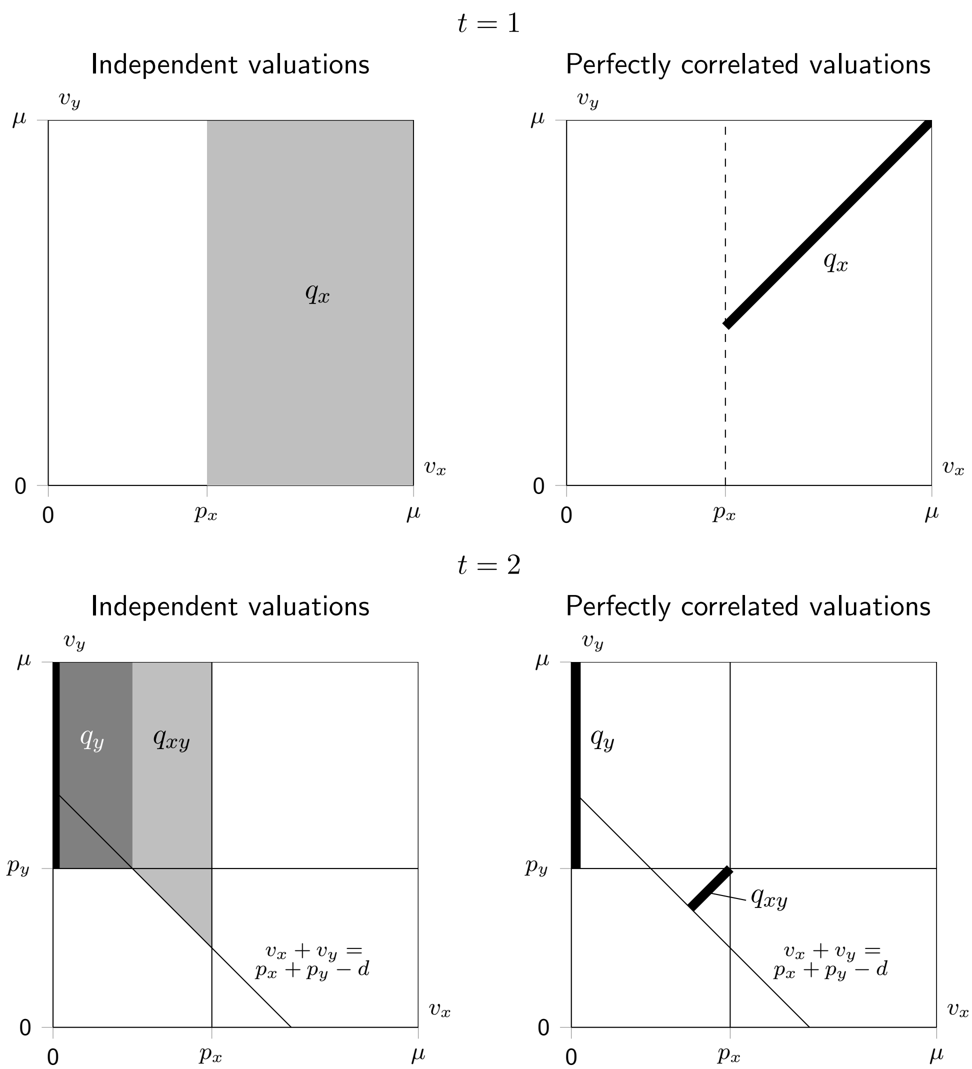
With only two types of buyers, the stylized model presented in Section 2 only allows for a qualitative analysis of the main effects of mixed bundling. In this section, we consider the continuous-type case and calibrate it with data from our sample of bundled movies. This exercise serves two purposes. First, to confirm the results from the theory section, in particular Proposition 2, which states that the seller's gain from bundling is strictly increasing in the extent to which valuations are correlated. Second, to compare the estimates from reduced-form regressions to those of a calibrated analytical model.

In Section 2, we assumed that consumer valuations can take two different values. Consider now the more realistic case where consumer valuations for a movie have a continuous distribution in $I\!R$. Specifically, we assume that v_i, the valuation for movie i, is uniformly distributed in $[0, \mu]$. Figure 9 illustrates the continuous case. As before, suppose there are two periods: at $t = 1$, only x is offered, whereas at $t = 2$ both y and the xy bundle are offered.

The top two panels correspond to $t = 1$, whereas the bottom two panels correspond to $t = 2$; the two left panels correspond to the case when v_1 and v_2 are independently distributed, whereas the right two panels correspond to the case when v_1 and v_2 are perfectly correlated.

Consider first the first period ($t = 1$) when valuations are independently distributed. If the seller sets p_x, then q_x is determined by the shaded are, basically all buyers whose valuations are greater than p_x. At $t = 2$, all of the mass in this gray area is concentrated along the v_y axis. In other words, all of the consumers who purchase x at $t = 1$ have zero valuation for x at $t = 2$; moreover, their valuation for y is uniformly distributed in $[0, \mu]$.

Figure 9
Bundling and sales with a continuum of types



At $t = 2$, the seller offers x and y for p_x and p_y (we assume $p_x = p_y$), and the xy bundle for $p_x + p_y - d$, where d is the bundling discount. Sales of y as a single are given by the area in dark gray (lower left panel of Figure 9) plus the mass along the y axis above p_y. In other words, there are two types of buyers of y at $t = 2$: high v_y, low v_x buyers who are making the first (and only) purchase; and high v_y, high v_x buyers who purchase x at $t = 1$ and y at $t = 2$. What these two groups have in common is that their valuations satisfy two conditions: $v_y > p_y$ (value for y greater than price) and $v_x + v_y - p_{xy} < v_y - p_y$ (surplus from purchasing bundle lower than surplus from purchasing y only). (The latter condition defines the boundary between the dark gray and the light gray areas in the lower left panel of Figure 9, whereas the former condition defines the lower boundary of the dark gray area.)

Still considering the case when valuations are independent and $t = 2$ (lower left panel), the consumers whose valuations fall in the light gray area purchase the bundle at $t = 2$. These are the consumers whose valuations satisfy two conditions: $v_x + v_y > p_{xy}$ (bundle valuation greater than bundle price) and $v_x + v_y - p_{xy} > v_y - p_y$ (surplus from purchasing bundle greater than surplus from purchasing y only). (The latter condition defines the boundary between the light gray and the dark gray areas; the former condition defines the southwest boundary of the light gray area; finally, the right-boundary of the light gray trapezoid is defined by the area of $t = 1$ buyers.)

Overall, the seller's revenue per potential consumer at $t = 2$ (when valuations are independent) is given by

$$R_0(d) \;=\; p\left(\left(\frac{\mu-p}{\mu}\right)^2 + \left(\frac{\mu-p}{\mu}\right)\left(\frac{p-d}{\mu}\right)\right) + (2\,p-d)\left(\left(\frac{d}{\mu}\right)\left(\frac{\mu-p}{\mu}\right) + \tfrac{1}{2}\left(\frac{d}{\mu}\right)^2\right)$$

where we drop the subscript from the price variables (as we assume $p_x = p_y$). The 0 subscript in R_0 corresponds to independent valuations. The first term on the right-hand side of the first row corresponds to the dark-gray area (including the mass segment along the y axis). The second term corresponds to the light-gray trapezoid (a rectangle plus a triangle). (We divide all measures by μ so as to obtain revenue per potential customer.)

Suppose now that $v_x = v_y$ for all consumers. This situation is depicted in the two right panels in Figure 9. As before, at $t = 1$ all consumers with $v_x > p_x$ purchase product x. At $t = 2$, all of these consumers have zero valuation for x and a valuation for y which is equal to their valuation for x. It follows that the mass segment of probability that was previously located along the main diagonal and above (p_x, p_x) is now located along the y axis and above $(0, p_x)$.

At $t = 2$, *all* of these consumers make a purchase of y. In fact, the conditions for this to be their optimal choice are satisfied: $v_y > p_y$ (value for y greater than price) and $v_x + v_y - p_{xy} < v_y - p_y$ (surplus from purchasing bundle lower than surplus from purchasing y only).

A second set of consumers, those whose valuations are greater than $p - d/2$ but lower than p, purchase the bundle. These are the consumers whose valuations satisfy two conditions: $v_x + v_y > p_{xy}$ (bundle valuation greater than bundle price) and $v_x + v_y - p_{xy} > v_y - p_y$ (surplus from purchasing bundle greater than surplus from purchasing y only). Overall, the seller's revenue per potential consumer at $t = 2$ (when valuations are perfectly correlated) is given by

$$R_1(d) \;=\; p\left(\frac{\mu-p}{\mu}\right) + (2\,p-d)\,\tfrac{1}{2}\left(\frac{d}{\mu}\right)$$

The 1 subscript in R_1 corresponds to perfectly correlated valuations.

Note that first-period revenues are the same with independent or perfectly correlated valuations. We thus focus on revenues during the second period, when revenues depend on ρ (the degree of correlation in valuations) and on d (the bundling discount). There are several ways one can model the intermediate case when valuations are positively correlated but not perfectly correlated. Similarly to the simpler model in Section 2, we consider a mixture between the independent and the perfectly correlated cases, where ρ is the weight placed on the perfect correlation cases. This leads to a revenue $R(d, \rho)$ given by

$$R(d, \rho) = (1 - \rho)\, R_0(d) + \rho\, R_1(d) \tag{2}$$

We are interested in two different questions:

- Does mixed bundling lead to an increase in revenues; and, if so, by how much? Formally, how does $R(d, \rho)$ change as we increase d from $d = 0$ to positive values.
- How does the gain from mixed bundling depend on the degree of correlation in valuations? Formally, how does $R(d, \rho)$ change as we increase ρ all the way to $\rho = 1$.

In order to obtain a quantitative answer to these questions, we calibrate our continuous type model by using data on the median values of the price and quantity time series. Specifically, from the data shown in Figure 2, we see that a substantial fraction of sales take place during the first year; and there is a price decline from the first six months to the second six months after release. From the data for the median time series, total sales (in units) during the first six months are given by 299,3757; and during the second six months by 56,426. As to average price, it drops from \$15.06 to \$13.54. Assuming that valuations are uniformly distributed, this implies the following two parameter restrictions:

$$\begin{aligned} 299,3757 &= \eta\,(\mu - 15.06)/\mu \\ 56,426 &= \eta\,(15.06 - 13.54)/\mu \end{aligned}$$

where η is the total number of potential consumers and μ the upper bound of the distribution of valuations. Solving we get

$$\begin{aligned} \eta &= 858.798 \\ \mu &= \$23.11 \end{aligned}$$

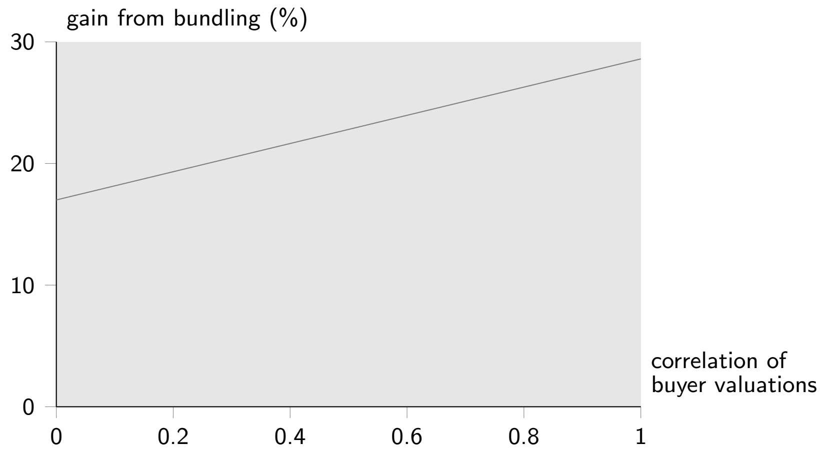
With these calibrated values at hand, we are now able to compute R, as given by (2), for different values of d and ρ. (We report the calibrated value of η for completeness sake. However, to the extent that we will report values on a per-potential-consumer basis, or on a percent variation basis, we will not make use of our estimate of η.)

Specifically, Figure 10 plots the gains from bundling as a function of the degree of correlation in valuations. Specifically, we compute the value of

$$G(\rho) \equiv \frac{R(5, \rho) - R(0, \rho)}{R(0, \rho)}$$

for ρ varying between 0 (independent valuation) and 1 (perfectly correlated valuations). The choice of $d = \$5$ is guided by Figure 6, which depicts the sample distribution of bundling discounts, with a median of about \$5.

Figure 10
Correlation of valuations and gains from bundling



As can be seen from Figure 10, we estimate the gains from correlation to be positive, that is, $R(5,\rho) > R(0,\rho)$. (Similar inequalities are obtained for other values of d, the bundling discount; more on this later). As such, our numerical computation confirms Proposition 1, the prediction from our simple theory model that the gains from mixed bundling are positive.

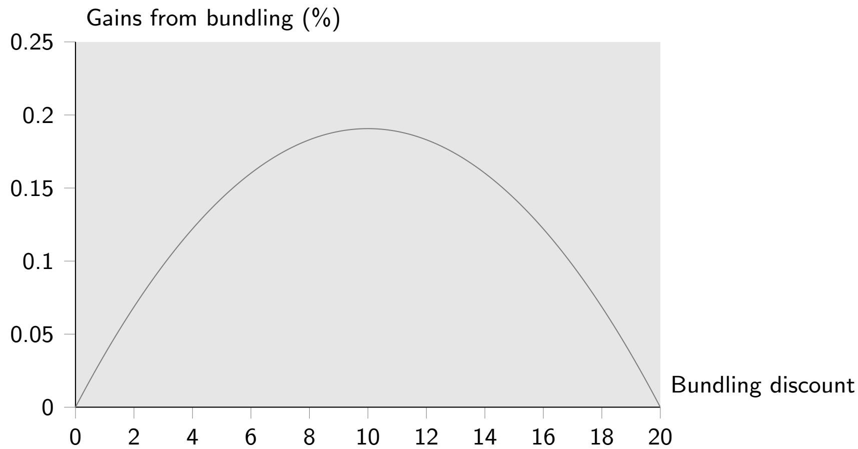
Consider now variations in the degree of correlation in valuations. For $\rho = 0$, we estimate a gain from bundling of about 17%, whereas for $\rho = 1$ this value is increased to about 28%. In other words, our estimates confirm Proposition 2, the prediction from our simple theory model that, around $\rho = 1$, the gains from mixed bundling are increasing in the degree of correlation in valuations.

If our numerical results are qualitatively consistent with the theory results from Section 2, we must also admit that, quantitatively speaking, the values from the calibrated model are lower than the estimated gains from the reduced-form regressions in Section 3. For example, the base estimates of the mixed-bundling dummy in the regressions show in Table 3 suggest a gain from mixed bundling in the order of 40%. Although this is greater than the 28% from the calibrated model (for $\rho = 1$), it is reassuring that we obtain values which are roughly of the same order of magnitude.

We do not have an obvious explanation for the discrepancy between these estimates of gains from bundling, but we can think of at least three candidates. First, as is well known the distribution of revenues in the movie industry is highly skewed, and this skew may affect our reduced-form estimates in ways that our calibrated model does not reflect. Second, we made some specific assumptions regarding the distribution of valuations. For example, if the distribution is uniform with a strictly positive lower bound then we get higher values of gains from bundling. Third, despite the long list of controls we use in our reduced-form regressions, it is possible that some unobserved factor explains both revenues and the decision to bundle.

All of these caveats notwithstanding, we believe the general gist of our theoretical, empirical and numerical evidence corroborates the central point of our paper: that there are substantial gains from mixed bundling and that these are increasing in the degree of

Figure 11
Bundling discount and gains from bundling



correlation in consumer valuations.

In the above calculations, we used $d = \$5$ as the base value for the bundling discount. However, as Figure 6 shows, there is a significant spread in the value of the d in our sample. This suggests a natural question: how do the gains from bundling depend on the d (keeping singles prices fixed)? Figure 11 answers this question. It shows that the gains from bundling are a quasi-concave function of the bundling discount, ranging from zero (no discount or discount of \$20) to about 20% (discount of about \$10).

Our estimated optimal discount (about \$10) is considerably greater than the median discount in our sample (about \$5). One possible justification for this gap is that several bundles are more than just a bundle: they included added value (e.g., "special edition"). Unfortunately, we do not have data on which bundles include value added and how much it is worth in the eyes of consumers, but to the extent that this happens often and corresponds to significant value, this can explain at least in part the gap between our \$10 number and the median value of \$5. We also note that, as documented by Figure 6, there is significant variation in the sample value of d, including in fact some negative values, which is consistent with the interpretation that some bundles include significant value added with respect to the sum of the values of the singles. Finally, different distributional assumptions lead to different optimal values of the bundling discount. For example, if the lower bound of the distribution of v is strictly positive then we obtain a lower value of the optimal bundle discount.

5. Discussion and concluding remarks

Searching on the Internet for commentary on the strategy of bundling, we came across a post by a "pricing strategy consultant" (an economics PhD) stating that

> Refer to any pricing book or ask any pricing expert when a company should implement a mixed bundling strategy and inevitably, you'll get the Adams and Yellen explanation. ...

> In my opinion, where the authors lose touch with reality is their important assumption of the key reason when to implement mixed bundling. The authors argue that mixed bundling should only be implemented when you have customers that have negatively correlated demands for products in a bundle. ...
>
> While the Adams and Yellen mixed bundling explanation may make sense theoretically, I don't think it's realistic in the real world.[11]

We argue that, in an intertemporal price discrimination context, the assumption that "customers have negatively correlated demands" is far from unrealistic; in fact, it is quite natural.

We provide three sources of corroborating evidence for our sequential-release interpretation. First, our simple, two-type theory model formalizes the qualitative point that sequential release of durable goods turns positively correlated valuations into negatively correlated valuations. Second, our reduced-form regressions — a difference analysis around the introduction of a bundle — suggest that mixed bundling leads to an increase in revenues and that this increase is greater the greater the similarity between the titles included in the bundle. Finally, our continuous-type model confirms the results from the two-type model and, once calibrated with mean sample values, implies revenue increases and gains from bundling that are broadly consistent with the reduced-form regressions.

We considered several alternative solutions to the "positive correlation" puzzle. One is that, although two different movies have positive correlation in characteristics, the correlation in valuations is actually negative.[12] A second one is that, as mentioned by Gandal et al. (2015) (who in turn refer to Johnson and Myatt (2006)), there are gains from bundling in a static framework even when valuations are positively correlated.[13] Finally a bundling offer may simply be a type of recommender system ("if you like *Conspiracy Theory* you are probably also interested in *The Pelican Brief*").[14] While these alternative narratives certainly have explanatory power in other contexts, we believe that, in the present context, our explanation makes more sense. For example, the market expansion effect considered by Gandal et al. (2015) and Johnson and Myatt (2006) requires that mixed bundling be sufficiently close to pure bundling, which, as Figure 7 shows, is far from being the case in our sample. Moreover, the Amazon-type recommender system explanation would normally be associated with a zero bundling discount (as is the case on Amazon), which, as Figure 6 shows, is far from being the case in our sample.

We believe the phenomenon we characterize in this paper has relevance beyond DVD sales. Many media products share several properties with the DVD industry. For example,

11. http://www.pricingforprofit.com/pricing-strategy-blog/interested-in-mixed-bundling-i-think-you-are-being.htm, visited October 14, 2016.
12. For example, consumers are interested in watching one and only one movie starring Julia Roberts. Therefore, if they buy *Pelican Brief* then the value of *Conspiracy Theory* is zero or substantially lower.
13. Their idea is that the Stigler (1963) model misses an important point, which they 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. (2015) observe (empirically), this is consistent with a revenue-increasing effect of bundling.
14. As often happens on Amazon's site, buyers are told that their chosen item is "Frequently bought together" with some other item together with a price offer for the bundle.

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, you can 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. Finally, as mentioned earlier, Derdenger and Kumar (2013) consider the case of bundling videogames with videogame consoles, an example that also fits our story.

Appendix

Proof of Proposition 1: Suppose that no sales occur at $t = 1$ and consider the seller's problem at $t = 2$. This is the "classic" bundling problem, that is, the two products are available at the same time. Since valuations are perfectly correlated, optimal prices are the same as if there were only one product on sale; in other words, bundling does not increase revenues. Since $\underline{u} > \alpha\,\overline{u}$, it is optimal for the seller to set $p_{x2} = p_{y2} = \underline{u}$.

Consider now the case when all buyers (high- and low-valuation ones) purchase product x at $t = 1$. Then, at $t = 2$, the problem is essentially the same as the one in the preceding paragraph, with the difference that only product y is sold. Optimal price is given by $p_{y2} = \underline{u}$. Moreover, bundling plays no role, since there is only demand for one product.

Finally, consider the case when high-valuation buyers purchased product x at $t = 1$.[15] If no bundling is offered at $t = 2$, then optimal prices absent bundling are given by $p_{x2} = p_{y2} = \underline{u}$ (by the same argument as in the preceding paragraphs). This yields the seller a $t = 2$ revenue of $(1-\alpha)\,\underline{u} + \underline{u}$: a fraction $1-\alpha$ buyers purchase product x (low valuation buyers) and all buyers purchase product y.

Alternatively, the seller may offer a bundle xy for a price $b = \underline{u} + \underline{u}$ to attract low-valuation buyers (who have not purchased product x at $t = 1$) and set $p_{y2} = \overline{u}$, a price targeted at high-valuation buyers. Since $\overline{u} < 2\,\underline{u}$, high-valuation buyers prefer to purchase product y rather than the bundle xy (having purchased x at $t = 1$, these buyers have no value for a second unit of x). Low-valuation buyers, in turn, purchase the bundle. This bundling strategy yields the seller a $t = 2$ revenue of $(1-\alpha)\,2\,\underline{u} + \alpha\,\overline{u}$: low-valuation buyers, a mass of $1-\alpha$, purchase the bundle for $2\,\underline{u}$; and high-valuation buyers, a mass α, purchase product y for $\overline{u}$. This is strictly greater than profit under no bundling, for $(1-\alpha)\,2\,\underline{u} + \alpha\,\overline{u} > (1-\alpha)\,\underline{u} + \underline{u}$ is equivalent to $\overline{u} > \underline{u}$.

Consider now pricing at $t = 1$ and suppose that the seller sets a price to attract high-valuation buyers. The latter correctly anticipate that, at $t = 2$, $p_{y2} = \overline{u}$ and $b = 2\,\underline{u}$. The highest price p_{x1} that high-valuation buyers are willing to pay for product x is therefore given by equality

$$(\overline{u} - p_{x1}) + \delta\,(\overline{u} - p_{y2}) = \delta\,(2\,\overline{u} - b)$$

which implies

$$p_{x1} = \overline{u} - 2\,\delta\,(\overline{u} - \underline{u})$$

Total seller profit is then given by

$$\pi_{B} = \alpha\left(\overline{u} - 2\,\delta\,(\overline{u} - \underline{u}) + \delta\,\overline{u}\right) + (1-\alpha)\,\delta\,2\,\underline{u} \tag{3}$$

Alternatively, the seller sets a higher p_{x1}, so that no buyer makes a purchase at $t = 1$. This leads to the subgame described in the first paragraph, which in turn corresponds to a total profit of

$$\pi_{N} = \delta\,2\,\underline{u} \tag{4}$$

Note that (3) can be rearranged as

$$\pi_{B} = \alpha\,\overline{u}\,(1-\delta) + \delta\,2\,\underline{u} \tag{5}$$

which is clearly higher than the value of π_{N} given by (4). We thus conclude that, at $t = 1$, high-valuation buyers purchase product x; and at $t = 2$ high-valuation buyers purchase

15. The case when only low-valuation buyers purchase x at $t = 1$ cannot be part of an equilibrium.

product y, whereas low-valuation buyers purchase the bundle. ■

Proof of Proposition 2: First notice that 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, notice that 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.

From the proof of Proposition 1, if $\rho = 1$ total profit is given by

$$\pi_B = \alpha\left(\overline{u} - 2\,\delta\,(\overline{u} - \underline{u}) + \delta\,\overline{u}\right) + (1-\alpha)\,\delta\,2\,\underline{u}$$

This corresponds to high-valuation buyers purchasing product x at $t = 1$ and product y at $t = 2$ (as singles); and low-valuation buyers purchasing the bundle at $t = 2$.

For ρ different from, but close to, 1, we must consider a fraction $\alpha\,(1-\alpha)\,(1-\rho)$ buyers with low valuation for product x buy high valuation for product y. These purchase the bundle xy at $t = 2$ but make no purchase at $t = 1$, just like buyers with low valuation for both products. In other words, compared to buyers with low valuation for both products, this change in valuations implies no change in revenue.

We must also consider a fraction $\alpha\,(1-\alpha)\,(1-\rho)$ buyers with high valuation for product x buy low valuation for product y. These purchase product x at $t = 1$ but make no purchase at $t = 2$ (for the valuation for product y is lower than the q or b). Compared to buyers with high valuation for both products, this implies a loss of $\alpha\,(1-\alpha)\,(1-\rho)$ (on a per-potential-consumer basis).

Finally, it is straightforward to check that the revenue loss is decreasing in ρ. ■

References

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." *The American Economic Review*, (2011), pp. 263–303.

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. "Ain't it 'Suite'? Bundling in the PC Office Software Market." Working Paper (2015).

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." *The 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.

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).