Mixed Bundling in Retail DVD Sales: Facts and Theories

Luís Cabral
New York University

Gabriel Natividad
Universidad de Piura

This draft: December 2017

Abstract. We examine the practice of bundling sequentially-released durable goods such as movie DVDs. We first present some reduced-form empirical evidence: most bundles consist of two different titles and are introduced soon after the second title release; bundles originate from the same distributor and consist of similar titles (user rating, box-office revenue, lead actor, etc.); finally, we estimate that the gains from bundling are greater the greater the similarity across titles included in the bundle.

To the extent that correlation of characteristics is related to correlation of valuations, the evidence seems at odds with the conventional wisdom that the gains from bundling are greater when valuations are negatively correlated. We propose an explanation based on the idea that, by the time a second DVD is released, high-valuation buyers have already purchased the first DVD. This implies that, even if from the beginning buyer valuations are positively correlated, by the time the second DVD is released they are negatively correlated, which in turn makes bundling a revenue-increasing strategy.

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 and the Munich Summer Institute for useful comments. Cristián Figueroa, Diego Zúñiga and Timothy Petliar provided excellent research assistance. The usual disclaimer applies.

1. Introduction

Many retail stores, such as Walmart, sell DVDs of previously released movies. In some cases, DVD titles are sold in bundles. Typically, the 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). The bundle is typically sold along the separate titles (mixed bundling). As theory would predict, the bundle price is typically lower than the sum of the separate titles' prices. For example, Universal Pictures' *The Scorpion King*, starring Dwayne Johnson, was released in 2002 for a price of about $18. 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*. While there were price variations across stores and from week to week, the mean price for the bundle was about $20, whereas the prices for *The Scorpion King* and *The Rundown* as singles were about $9 and $19.5, respectively.

In several ways, the retail DVD sale market is similar to many other markets where mixed bundling is a common marketing strategy. However, 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.

In this paper, we examine the practice of bundling sequentially-released durable goods such as movie DVDs. We first present some reduced-form empirical evidence regarding the practice of mixed bundling. Among other facts, we establish that most bundles consist of two different titles and are introduced soon after the second title release. Moreover, bundles originate from the same studio and consist of similar titles (user rating, box-office revenue, lead actor, etc.). Next, we attempt to estimate the effects of bundling. We estimate that, upon the introduction of a bundle, individual prices increase and individual sales decrease, whereas overall revenues increase by about 40%. We also estimate that the gains from bundling are greater the greater the similarity across titles included in the bundle.

Next we consider theories that account for the empirical results, in particular the positive correlation between gains from bundling and similarity of titles included in a bundle. At face value, this results seems in contradiction with the common wisdom that the gains from bundling are greater when buyer valuations are negatively correlated across products. One set of theories argues that, despite the positive correlation in product characteristics, valuations are actually negatively correlated. Another theory is based on the idea that, if mixed bundling is close to pure bundling (i.e., sales of singles are very small) then the gains from bundling are increasing in the degree of correlation of buyer valuations.

Unlike these theories, which have the common characteristic of being static models, our preferred theory is dynamic and fits particularly well two features of DVDs (and many other markets): durability and sequential release. It is well known that bundling and sequential release of versioned products are two alternative ways to price discriminate. We show that the combination of bundling and sequential releases creates a powerful additional tool to price discriminate. The idea is that, by the time a second DVD is released, high-valuation buyers have already purchased the first DVD. This implies that, even if from the beginning buyer valuations are positively correlated, by the time the second DVD is released they are negatively correlated, which in turn makes bundling a revenue-increasing strategy.

We also consider an additional theory, one that is based on limited awareness of supply (Hendricks and Sorensen, 2009). In this case, bundling functions as a sort of "recommender system" used by online sellers such as Amazon. The value of bundling is quite clear: if a consumer is willing to pay p for one product then such consumer is also willing to pay p for a similar product. We argue that the data is more consistent with a sorting theory than with a limited awareness theory.

■ **Related literature.** The theoretical literature on bundling dates back to Stigler (1963), who showed with a simple example that bundling can be a profitable strategy. Adams and Yellen (1976) and Schmalensee (1984), among many others, expanded on this theme, presenting a more systematic approach to the conditions under which bundling is profitable. The most systematic approach to this problem, in particular relating the gains from bundling to the correlation in buyer valuations, is probably given by Chen and Riordan (2013). They show 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.*

Anecdotal evidence, and the more systematic data we present in Section 2, shows that DVDs that are bundled are typically very similar, which in turn suggests valuations are highly correlated. We propose a way to reconcile this strand of theory with the evidence, namely to explicitly account for sequentiality in sales.

Empirically, Gandal et al. (2015) directly addresses the issue of correlation of preferences. They estimate a discrete-choice model of software demand and apply it to 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 (video-game 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. 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]

Figures 1 and 2 provide some evidence on the dynamics of unit sales and prices. In each case, we represent the median as well as the 5th and 95th percentiles of the distributions of q and p for each week since release. In the case of q, the 10th percentile is so close to zero that cannot be distinguished from the x axis.

Regarding sales (in 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, just as unit sales we notice a decline over time, though at a much lower rate. The median price starts at about \$15, and after 1.5 years stabilizes at about \$10. As the 5h and 95th percentile curves show, there is considerable variation. Moreover, as we show below, the distribution of price is typically multi-modal, which results from the fact that a considerable portion of the probability mass is concentrated around even or near-even numbers (e.g. \$15, \$20, etc).

■ **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 always offered in a mixed-bundling regime, that

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 include 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.
5. We determine an item is a bundle when its name includes the names of different feature films.

Figure 1

Sales (quantity) over time

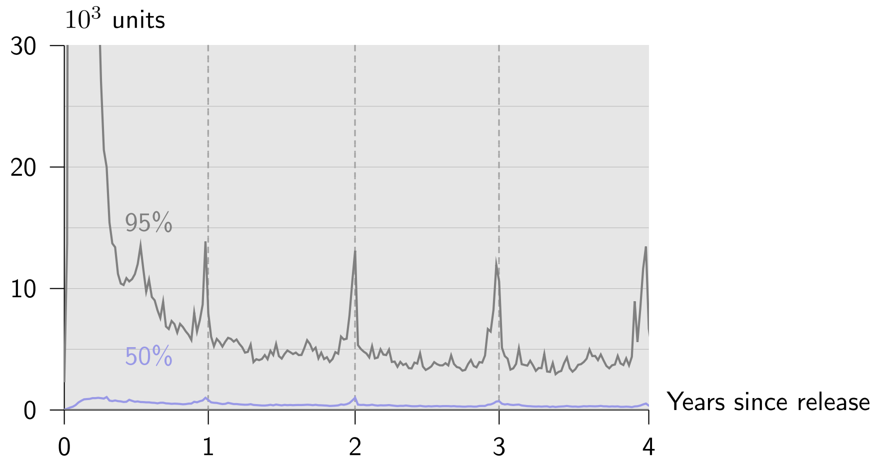


Figure 2

Prices over time

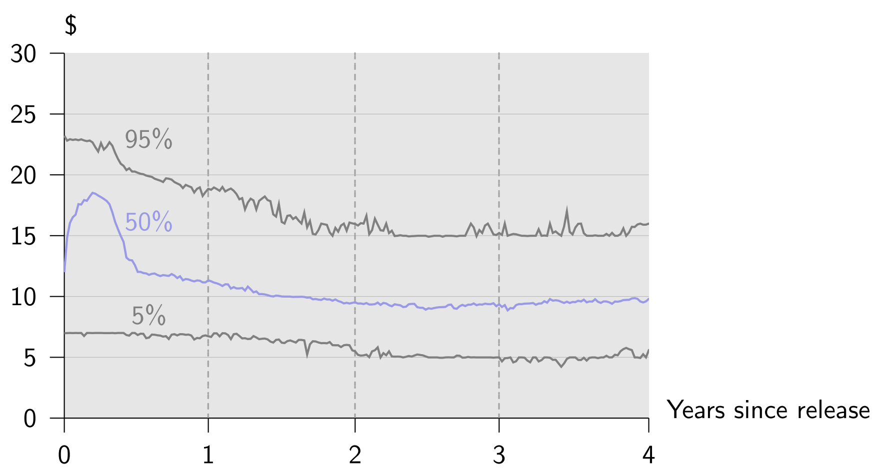
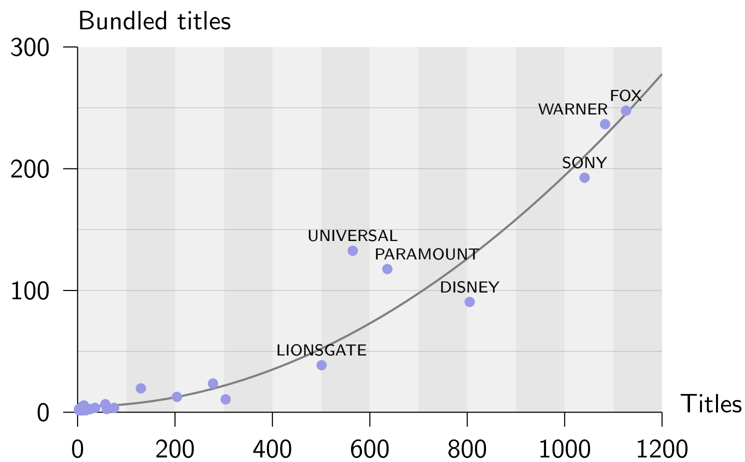


Table 1
Summary statistics for bundles

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



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 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.

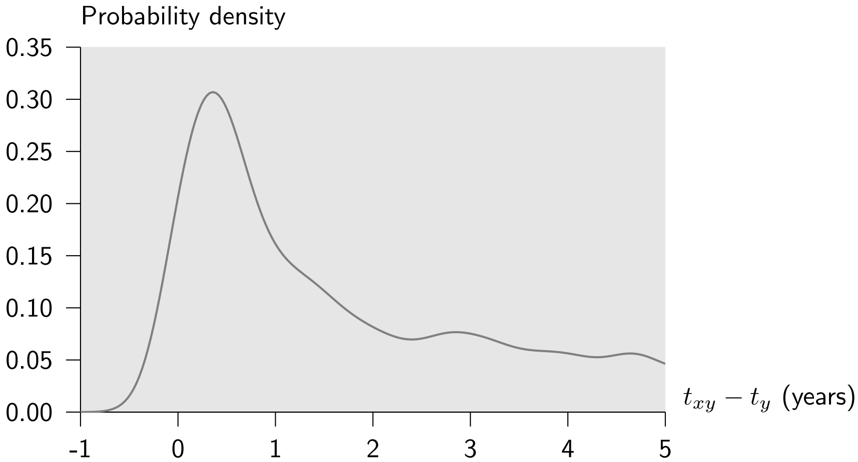
The message is clear: bundles are not random pairings. Rather, bundles disproportionately combine DVDs of a similar genre, language, MPAA rating, movies with the same

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

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



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 terms of user rating. They do, 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.[8]

■ **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,

$$p_x + p_y - p_{xy}$$

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

8. There is also a small number of titles when $t_{xy} < t_y$, that is, the second title is first released in a bundle with the first title.

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

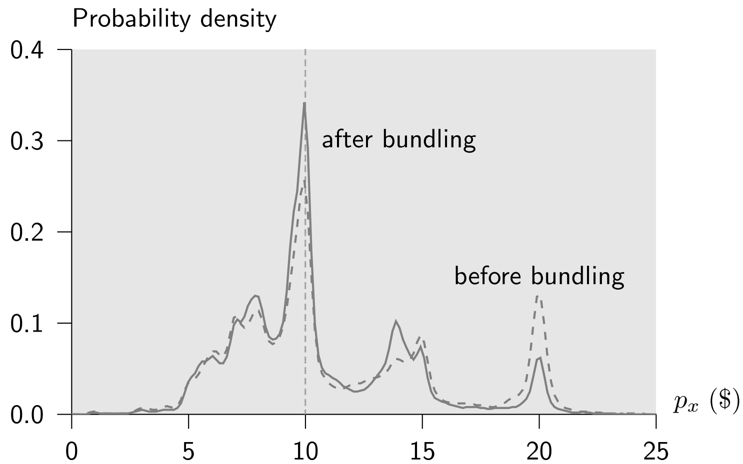


Figure 6
Bundling discount

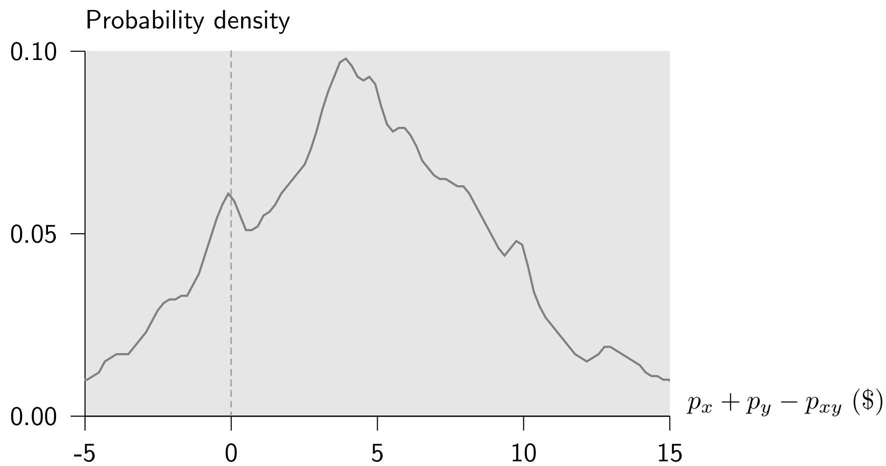
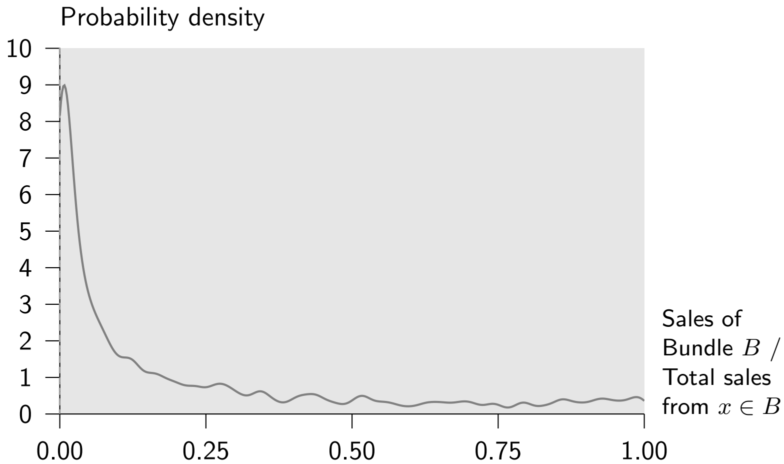


Figure 7
Bundle sales and total sales



across stores. This could therefore be an artifice of aggregation.[9]

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 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 that 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 shows a distribution that is approximately uniform 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 we are before a case 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.

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

Most of these facts are probably not surprising. What seems surprising — at least to us — is 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."[10]

■ **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 does 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 Figures 1 and 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 first strategy to take these problems into account is to (a) include calendar time and movie 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.[11] 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 on 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%.

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

11. 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

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 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

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

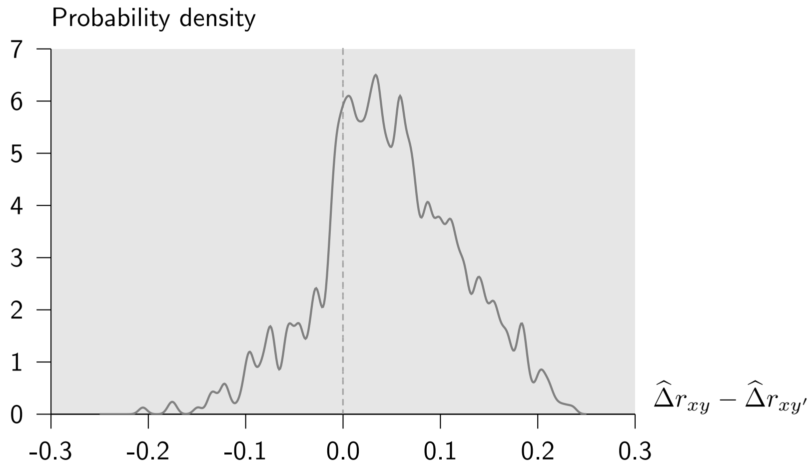
of regressions suggests that, for bundles that were relevant for total unit sales, a bundle is associated with an increase in total unit sales (almost a doubling of total unit sales, an increase of 43.1+54.4=97.5%), whereas the sales of singles drops by about 8.6%.

■ **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 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 redo 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

Figure 8
Bundles and pseudo-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 checks on these results. First, our analysis is done at the movie-title level, something we do by pro-rating bundle revenues and unit sales to the constituting bundle component titles. In the process, we treat symmetrically all of bundle titles. 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 pro-rate 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 2, 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 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 develop a stylized theoretical model that is consistent with these various pieces of evidence. We then (Section 4) calibrate a structural model that allows us to run a series of counterfactual exercises.

3. Theory

In the previous section, we uncovered a series of stylized facts regarding mixed bundling in the retail DVD sales market. We showed that distributors tend to create bundles that combine similar movies, namely in terms of the identity of the lead actors. We showed that mixed bundling increases seller revenues and that the revenue increase is increasing in the degree of similarity between the titles included in the bundle. Regarding the effects of bundling, our analysis also suggests a small or negative effect on unit sales of singles, a small increase in the price of singles, and a positive bundling discount.

To the extent that similarity of characteristics is associated with correlated preferences, these results suggest that the gains from mixed bundling are greater the greater the correlation in valuations across the goods included in the bundle. This stylized fact seems at odds with the the conventional theory of mixed bundling. For example, summarizing and extending a long literature that dates back to Stigler (1963), Chen and Riordan (2013) state 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.

In this section, we propose a theory that fits the data reasonably well, in particular the observation that the gains from mixed bundling are increasing in the degree of similarity between movies.[12] Our theory is based on a dynamic model which fits particularly well two features of DVDs (and many other markets): durability and sequential release. Given the centrality of this theory, and its novelty with respect to standard bundling theory, we present a formal model and formal propositions that establish the relation between theory and data. This subsection is structure 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. By contrast, we show that, in a sequential-release context, bundling as a positive effect on revenues. Second, we show that the gains from bundling are increasing in the degree of similarity across products, in a way that fits the empirical evidence.

■ **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 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 make an important assumption regarding bundling, namely that the seller cannot commit at $t = 1$ not to offer a bundle at $t = 2$. Later we consider the alternative assumption whereby the seller can commit not to bundle.

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

12. In Section 5 we present various alternative theories.

Figure 9
Imperfectly correlated valuations

	H	L
H	$\alpha\left(\alpha(1-\rho)+\rho\right)$	$\alpha(1-\alpha)(1-\rho)$
L	$\alpha(1-\alpha)(1-\rho)$	$(1-\alpha)\left((1-\alpha)(1-\rho)+\rho\right)$

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 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 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 9 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 buyer: 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 Bro's *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.

Figure 4, introduced earlier, 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.

Strictly speaking, examples like *The Pelican Brief*/*Conspiracy Theory* violate 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.

4. Structural model

With only two periods and two types of buyer, the stylized model presented in the previous section only allows for a qualitative analysis of the main effects of mixed bundling. In this section, we develop a dynamic structural model of consumer demand with a multiple periods and buyer types. The advantage of this approach is to allow for a more precise understanding of the qualitative effects of bundling, as well as comparative statics with respect to crucial variables such as release date, preference correlation and bundling discount.

Suppose that time t is discreet, specifically with each period corresponding to one week. For the purpose of simulations, we consider a five-year horizon, $t = 0, ..., 260$. Consumer valuations for each of a pair of movies, (v_1, v_2), are distributed according to a symmetric joint cdf $F(v_1, v_2)$ with density $f(v_1, v_2)$. Given that we assume a symmetric distribution, with some abuse of notation we denote by $F(v_i)$ and $f(v_i)$ the marginal cdf and pdf.

In our numerical simulations we consider a discrete grid of consumer valuations, in which case $f(v_1, v_2)$ and $f(v_i)$ correspond to the measure of consumers with valuations in a given cell of valuation levels. Moreover, we will assume that the marginal distributions are uniform, $U[0, \psi]$.[13] It follows that the marginal density (the same for both valuations) is given by

$$f(v_i) = 1/\psi$$

13. A more realistic assumption is that a share γ of the population has zero valuation, whereas the remaining $1 - \gamma$ have uniformly distributed valuations. However, we cannot separately identify total population η and the fraction $1 - \gamma$ that has non-zero valuation.

In order to model correlation, we first consider two extreme cases: independent valuations, i.e., $f(v_1, v_2)$ is simply the product of the marginals; and perfectly correlated valuations, i.e., $f(v, v) = 1/\psi$ and $f(v_1, v_2) = 0$ if $v_1 \neq v_2$. More generally, we consider the intermediate case when the density $f(v_1, v_2)$ is a mixture of the zero and 1 correlation cases, with $\rho \in [0, 1]$ being the weight placed on perfect correlation; ρ is therefore a measure of the correlation of valuations.

In order to be consistent with the stylized pattern of consumer demand over time, we assume that valuations decrease with movie age. Specifically, movie i's valuation at time t is given by v_i times an age factor α_t, which is given by

$$\alpha_t = \begin{cases} 1 & \text{if } \; t \leq \tau \\ \omega + (1 - \omega)\,\frac{\tau}{t} & \text{if } \; t > \tau \end{cases} \tag{2}$$

where $0 < \omega < 1$. Note that $\alpha_0 = 1$; and that, as $t \to \infty$, $\alpha_t \to \omega < 1$. We assume the good in question is a durable good and that valuations correspond to at most one unit. This implies that, once a consumer makes a purchase, her valuation for the good drops to zero.

Let η be population size, which we assume is constant over time. Each week (at the beginning of the week), a fraction ν of the population, a measure $\nu\,\eta$, dies; and an equal measure $\nu\,\eta$ is born. As in the perpetual-youth model (Blanchard, 1985), all consumers, regardless of valuation or purchase history, are equally likely to die.[14]

Each week, after deaths and births have taken place, a living consumer is active with probability ϕ. By "active" we mean that the consumer consciously decides whether or not to make a purchase. We assume active consumers behave myopically, that is, choose to purchase a movie or a bundle of movies based on their current utility. Considering the low probability that a movie is bundled; and considering that prices decline at a relatively low rate, we believe this is a good approximation for consumer behavior.

Let $g_t(v_1, v_2)$, $v_i > 0$, be the measure of living consumers with valuations (v_1, v_2) who have not purchased the product by the beginning of period t (after birth and death events have taken place). At $t = 0$,

$$g_0(v_1, v_2) = \eta\, f(v_1, v_2)$$

For $t > 0$, $g_t(v_1, v_2)$ is updated based birth, death and purchase decision. Below we go into this in greater detail.

■ **No bundling.** Consider first the case of no bundling. (For about 90% of the DVD titles in our sample, this is the relevant case.) This case provides a helpful stepping stone for the more general case. It also allows us a simple strategy for calibrating the model.

Given that no bundling ever takes place, and considering our consumer behavior assumption, we can simply work with marginal densities for each valuation level. We also denote by v the valuation in question. For each time t, we consider two possibilities. If $p_t < \alpha_t\, v$, then an v-valuation active consumer makes a purchase. Below we denote by v_t the effective value at time t, that is, $v_t \equiv \alpha_t\, v$. It follows that $g_t(v)$ is updated as follows:

$$g_t(v) = (1 - \phi)\,(1 - \nu)\, g_{t-1}(v) + \nu\,\eta\, f(v)$$

14. Whereas time is treated as discrete, valuations are treated as a continuous variable; and the number of consumers is considered large enough that we can indistinctly work with numbers and measures. In other words, the probability and the frequency approaches are equivalent.

Figure 10
Model calibration

Param.	Value	Interpretation
η	2000	2 million buyers have non-zero valuation
ϕ	.03	live consumer has 3% probability making decision
ν	.005	each week half of a percent of the population "dies"
ψ	50	highest valuation for a DVD (in $)
τ	3	obsolescence begins after 3 weeks
ω	.2	valuations converge to 22% of the original valuation

Intuitively, of the $g_{t-1}(v)$ consumers present in the previous period, a fraction ϕ were active and made a purchase, thus leaving the remaining fraction $(1-\phi)$ to be carried into the current period. Of these, a fraction ν die at the beginning of the period, thus leaving a fraction $(1-\nu)$ of survivors. Finally, a measure $\nu\,\eta$ of consumers is born; and of these, a fraction $f(v)$ is endowed with valuation v.

Suppose instead that $p_t > \alpha_t\,v$. Then a v-valuation consumer does not make a purchase and $g_t(v)$ is updated purely based on consumer birth and death:

$$g_t(v) = (1-\nu)\,g_{t-1}(v) + \nu\,\eta\,f(v)$$

Sales at time t are given by

$$q_t = \phi \int_{p_t/\alpha_t}^{\infty} g_t(v)\,dv$$

Finally, we take care of the special case $v=0$ by considering a mass point g_t° that evolves according to the equation

$$g_{t+1}^\circ = (1-\nu)\,g_t^\circ + q_t$$

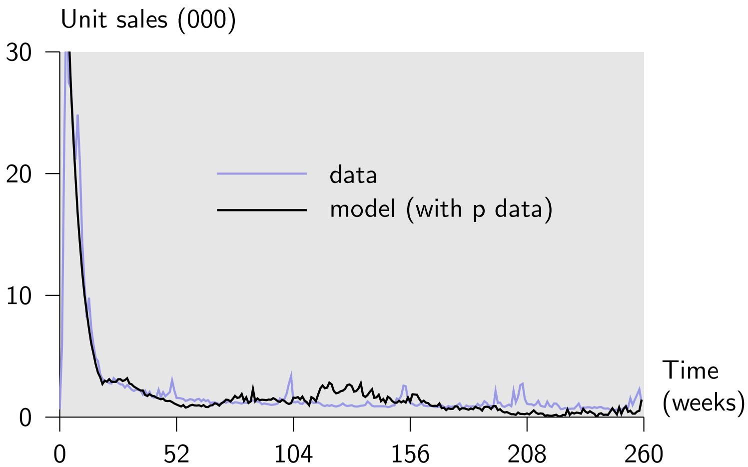
In words, a fraction $(1-\nu)$ of consumers with exactly zero valuation die, whereas all consumers who made a purchase, q_t, become zero-valuation consumers.

■ **Calibration.** The no-bundling case allows us to calibrate the model parameters (η, ϕ, η, ν, ψ, τ, ω) so that, given mean prices at time t (cf Figure 2), we reproduce median sales values at different periods of a movie's sales path (cf Figure 1).

Table 10 shows the estimated parameter values. In words, $\eta = 2000$ means that there are a total of 2 million interested buyers (that is, with non-zero valuations); $\phi = .03$ means that a live consumer has a 3% probability of deciding whether to purchase the DVD in a given week; $\nu = .005$ means that each week half of a percent of the population dies (meaning, leaves the market) and is replaced by a an equal number of consumers; $\psi = 50$ means that the highest valuation for a DVD is \$50; and $\omega = .2$ means that, in the limit as time goes by, valuations for the movie converge to 20% of the original valuation.

Figure 11 shows both the data and the model prediction (which, as we mentioned before, uses median prices as the values of p_t). Although we have six parameters to work with, we force ourselves to accept a specific dynamic structure. In others words, a six-parameter polynomial would do a better job at approximating the time series for q, but it would have no underlying economic structure.

Figure 11
Sales time path



One of the failures in calibration is that the model exaggerates the consumers' sensitivity to price variations: during the third year mean prices are lower but sales do not increase accordingly.

■ **Mixed bundling.** Having calibrated the model, we can now proceed to use the model to study the effects of mixed bundling. Suppose that movie 1 is released at time $t = 0$ and movie 2 is released at time $t = 156$ (3 years after the first release). Finally, suppose that a bundle is released at $t_b \geq 156$.

For $t < 156$, the model is solved as in the case when there is no bundling. For $156 < t < t_b$, the same solution also applies, with the difference that we need to keep track of both valuation levels.

The critical difference corresponds to the case when $t \geq t_b$. Now, for each pair of valuations (v_1, v_2), we must consider four possibilities

- $v_{1t} < p_{1t}$, $v_{2t} < p_{2t}$ and $v_{1t} + v_{2t} < p_{bt}$, in which case no purchase takes place
- $v_{1t} > p_{1t}$ and $v_{2t} < p_{2t}$, in which case the active consumer purchases title 1 only
- $v_{1t} < p_{1t}$ and $v_{2t} > p_{2t}$, in which case the active consumer purchases title 2 only
- $v_{1t} > p_{1t}$, $v_{2t} > p_{2t}$ and $v_{1t} + v_{2t} > p_{bt}$, in which case an active consumer purchases the bundle

As before, and consistently with the assumption that consumers have a preference for at most one unit of each movie, we update consumer valuations after each purchase: if a consumer buys movie i then its valuation switches to zero.

■ **Preference correlation and gains from bundling.** Having calibrated and model and computed the base-case equilibrium, we now proceed to perform a counterfactual that addresses the central issue considered in the "toy" model presented at the beginning of the section. Proposition 2 states that the gain from bundling is increasing in the degree of correlation of valuations, which we model in a simple two-type model. Does this property extent to a more realistic model with a continuum of valuations?

Figure 12
Correlation of valuations and gains from bundling

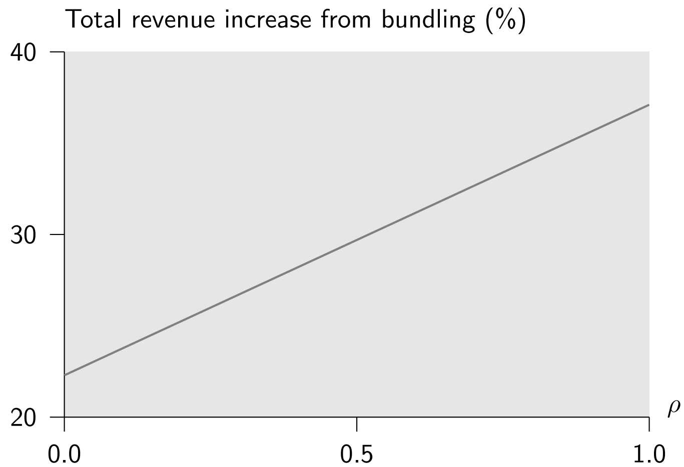


Figure 12 plots the gains from mixed bundling as a function of the degree of correlation in valuations. As mentioned earlier, we model correlation is a very simple way: we consider the cases of independent valuations $\rho = 0$ and perfect correlation $\rho = 1$ and then consider the more general case when the joint density of valuations is a mixture of the $\rho = 0$ and $\rho = 1$ cases. For each convex combination of this mixture we solve the model with its base-case parameter values and for values of p_x, p_y, p_{xy} equal to the sample mean (as in the above calibration).

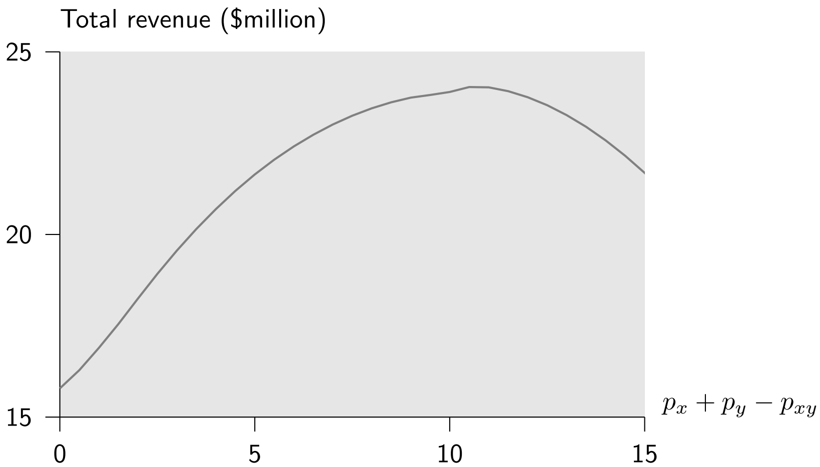
Consistently with Proposition 2, the gains from bundling are increasing in the degree of correlation, in an approximately linear fashion, from about 23% when $\rho = 0$ to about 38% when $\rho = 1$. Interestingly, the values are, broadly speaking, of the same order of magnitude (if a little smaller) as the estimates from the previous section.

One question which may be asked regarding our results is: if the gains from bundling are so high, how come only about one tenth of the movies are included in a bundle? Figure 12 also helps answering this question: The gains from bundling are considerably lower as the degree of correlation in valuations decreases. If we assume that there is a cost of bundling, which seems reasonable, then only a fraction of all possible bundles lead to a revenue increase that compensates for the bundling cost. Although we estimate a relatively high gain from bundling, this is the *average* gain from bundling, average among all actual bundles. Given the results in Figure 12, we expect the *marginal* gain from bundling to be considerably lower than the average. In other words, we do not need to assume a very high cost of bundling to justify an observed policy of bundling whenever there is a net positive gain from doing so.[15]

■ **Bundling discount and gains from bundling.** In our base-case simulation we set the bundling discount at the sample median (rounded to the next dollar): \$5; that is, we assumed that $p_x + p_y - p_{xy}$. A natural exercise to perform is to recalculate the model by changing p_{xy} while keeping p_x, p_y and all model parameters constant.

15. In order to test this interpretation, we would need to know the mapping from commonality of characteristics to correlation of valuations, which we don't.

Figure 13
Bundling discount and gains from bundling



How do the gains from bundling depend on the bundling discount? Figure 13 shows that the gains from bundling are a quasi-concave function of the bundling discount. If the bundling discount is zero, then the only positive effect of bundling is to increase awareness of existing movies; that is, some consumers may not be aware of the existence of titles i or j but become aware of the (i, j) offering. For a very high bundling discount, the seller effectively gives away the bundle, which in turn implies a zero or even negative contribution to revenues. For intermediate values of the bundling discount, the sorting effect we considered before kicks in, leading to higher revenues.

Figure 13 an optimal bundling discount of \$10.5. Considering that the mean price of a single title is about \$10 (cf Figure 2), Figure 13 suggests the optimal policy is to offer two for the price of one (at which point a consumer who already owns one of the titles is indifferent between buying a single and buying a bundle).

Our estimate of an optimal \$10.5 bundling discount is subject to two caveats. First, as we will show later the optimal bundling discount is sensitive to some of the model parameter values, including the distribution of consumer valuations: if we change the highest valuation from \$50 to \$100, then the estimated optimal discount changes from \$10.5 to \$4.5, which is the median value in our sample. Second, anecdotal evidence suggests that a good number of bundles are more than just a bundle: they correspond to "special editions" (e.g., of a movie and its sequel) which include extras. The value of these extras may explain part of the gap between the estimated optimal discount and the observed bundling discount.

■ **Sensitivity analysis.** In order to ascertain the sensitivity of our results to changes in parameter values, we perform alternative simulations where, for each parameter, we consider alternative values by doubling or halving the original value.

As expected, changes in η only have scaling effects: revenues from bundling and not bundling increase in equal amounts, and so the percent gain from bundling remains the same — and so does the estimated optimal bundling discount.

Changes in the frequency of purchase seem to have little effect on the optimal bundling discount. They do, however, on the estimated gain from mixed bundling: a higher ϕ leads

Table 6
Sensitivity analysis with respect to parameter values (holding $\rho = 1$)

Par	Description	Base	Alt	$\Delta\log(r)$	b^*
η	# buyers with non-zero valuation (10^3)	2000	4000	37.39	10.5
		2000	1000	37.39	10.5
ϕ	Probability making decision in given week	0.03	0.06	26.52	10.5
		0.03	0.015	47.64	10.5
ν	Consumer weekly hazard rate	0.005	0.01	43.08	10.5
		0.005	0.0025	32.75	11.0
ψ	Highest DVD valuation ($)	50	100	1.40	4.5
		50	25	8.65	16.0
τ	First week of declining valuation	3	6	15.64	9.0
		3	1.5	101.51	12.0
ω	Asymptotic valuation / initial valuation	0.2	0.4	2.90	7.0
		0.2	0.1	7.48	14.5
Base-case values:				37.39	10.5

to a lower gains from mixed bundling.

Changes in the consumer birth and death rate seem to have a relatively small impact on the optimal bundling discount or the gains from bundling. By contrast — and perhaps not surprisingly — the parameters governing the distribution of buyer valuations, ψ, τ, ω, seem to have a significant impact. Generally speaking, we find that parameters indicating higher consumer valuations imply lower gains from bundling; and a lower bundling discount.

5. Alternative theories

In the previous two sections we developed and applied a theory of bundling as a dynamic sorting strategy in the context of sequentially-released durable goods. In this section we consider alternative theories which are consistent with at least some of the stylized facts considered in Section 2.

5.1. Negative correlation of valuations

One possible explanation for the stylized fact that the gains from bundling are greater the greater the similarity in movie titles is that valuations are actually negatively correlated, which in turn implies gains from mixed bundling (Stigler, 1963; Adams and Yellen, 1976; Schmalensee, 1984; Chen and Riordan, 2013).

There are at least two possible narratives for negative correlation in valuations. First, suppose that some consumers are discriminating movie watchers, interested in one specific title but not in other titles (even in other titles that are very similar); whereas other consumers are equally interested in movies of a specific type (i.e., specific genre, featuring a specific lead actor, etc).

Such preference structure leads to negative correlation in valuations: within a given bundle, "discriminating" consumers have a very high valuation for one item and a very low valuation for the other; whereas non-discriminating consumers have the same (intermediate-level) valuation for both items. Note that the latter is only true if the two items are very similar; were the two items' characteristics very different, it would no longer be the case that non-discriminating buyers had the same valuation for both titles, and preference correlation across the two items would be less negative. Thus, this story is consistent with our central stylized fact.

A second narrative for negative correlation of preferences is that high-valuation consumers are only interested in watching one movie; whereas low-valuation consumers have constant valuation for each movie (i.e., they are "binge watchers"). Similarly to the previous story, this interpretation leads to a negative correlation in preferences: buyers who value one item very highly do not value the other item. Similarly to the previous story, the model is consistent with the revenue-increasing effect of mixed bundling.

5.2. Positive correlation in valuations

Gandal et al. (2015), who in turn refer to Johnson and Myatt (2006), make the point that the received wisdom regarding bundling — namely that it requires negative correlation in valuation — misses an important point, which they refer to as the market-expansion effect of bundling. 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.

Finally, what is true for pure bundling is also true (by continuity) for a situation that is close to pure bundling, that is, when most buyers either buy the bundle or nothing.

This view of bundling seems to have traction in the market for desktop computer software: consumers who have a high valuation for a word processor also have a high valuation for a spreadsheet software; and most consumers who buy software buy a bundle. Consistently with the theory, an estimated structural model suggests that revenues increase with bundling (mixed bundling which, in practice, is close to pure bundling).

An alternative narrative that is consistent with a static model of bundling and positive correlation of valuations views bundling as a form of non-linear pricing (quantity discount). The key assumption is that, regardless of type, consumers have a high valuation for one unit (one title) and a lower valuation for a second unit.

The quantity-discount narrative has one radically different prediction with respect to our dynamic sorting theory: whereas in our preferred theory bundles are purchased by low-valuation consumers, in the quantity-discount story bundles are purchased by high-valuation consumers.

5.3. Bundling as a recommender system

Most Amazon buyers are familiar with the experience: upon finding a product they searched, and its respective price, the system offers them an alternative deal: instead of buying item x for p_x, Amazon will sell x and y for price p_{xy}. Moreover, the price quoted for the bundle is typically the sum of the individual prices, that is, $p_{xy} = p_x + p_y$.

We now suggest a very simple theory that is consistent with this practice. As Hendricks and Sorensen (2009) have shown, media products such as CDs and DVDs add up to a very large number of different products. Consumers cannot be expected to be aware of all of them, or to have an accurate estimate of their value.

Hendricks and Sorensen (2009) show — theoretically and empirically — that, in this context, the purchase of one CD by band x may lead to an increase in the demand for a previous CD by the same band.[16] Similarly, the Amazon bundling strategy follows a similar principle: if the consumer is willing to pay p_x for x, then such consumer is also likely willing to pay $p_y = p_x$ for a product y that is similar to x, that is, a product whose valuation is highly correlated with that of x.

This setting naturally leads to a theory of bundling: to the extent that consumer attention and the seller's ability to produce a good mach are limited, bundling allows the seller to extend a good match from one sale to two sales. This theory is also consistent with the gains from bundling being increasing in the degree of similarity between the products included in the bundle: ultimately, if the products are very similar and buyer utility additive, then bundling would lead to a doubling of seller revenues.

However, we note that an Amazon-type model of bundling implies — as is the case with Amazon — a zero bundling discount. By contrast, our data clearly suggest a positive bundling discount.

6. Conclusion

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.[17]

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

Admittedly, the application we considered — DVD retail sales — is rather small in terms of economic relevance — and declining. However, we believe the phenomenon we characterize has relevance beyond DVD sales. 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, you can

16. Cabral and Natividad (2016) explore a similar backward spillover effect in the context of movie DVDs.

17. http://www.pricingforprofit.com/pricing-strategy-blog/interested-in-mixed-bundling-i-think-you-are-being.htm, visited October 14, 2016.

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 video games with video-game 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 $t = 1$.[18] 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

18. 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 lower 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 $\alpha\,(1-\alpha)\,(1-\rho)\,q$.

Finally, it is straightforward to check that the revenue loss is decreaing in q. ■

References

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

Blanchard, O.J. "Debt, Deficits, and Finite Horizons." *Journal of Political Economy*, Vol. 93 (1985), pp. 223–247.

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 Economics Review*, Vol. 97 (2007), pp. 713–744.

Hendricks, K. and Sorensen, A. "Information and the Skewness of Music Sales." *Journal of Political Economy*, Vol. 117 (2009), pp. 324–369.

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.

Schmalensee, R. "Gaussian Demand and Commodity Bundling." *Journal of Business*, (1984), pp. S211–S230.

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