CATÓLICA LISBON
BUSINESS & ECONOMICS

Analysis of sequential distribution on Video-on-Demand platforms

Laetitia Römer

Dissertation written under the supervision of
Miguel Godinho de Matos

Dissertation submitted in partial fulfilment of requirements for the
MSc in International Management,
at the Universidade Católica Portuguesa,
February 21, 2021.

Analysis of sequential distribution on Video-on-Demand platforms

Laetitia Römer *

February 21, 2021

Abstract

The sequential distribution of titles throughout different channels is a prevalent profit-maximising strategy in the movie industry. However, cross-channel effects can occur where a product's distribution in one channel affects sales in another channel. I analyse how the availability of a version of a movie in a digital channel impacts sales in another digital channel in the context of a Video-on-Demand platform offering titles to lease and to buy. I use secondary data from a Video-on-Demand platform, exploiting the sequential entry and exit of a movie's versions. I find that for movies available for more than one year before being offered as two versions, adding the second version increases the pre-existing version's sales. Contrarily, if movies are available for less than one year before being distributed as two versions, sales of the pre-existing version decrease. Thus, my results suggest that depending on the delay between introducing the version to lease and to buy adding the version to buy either has a positive information spillover or a cannibalisation effect.

Keywords: Versioning, Digital channel, Cross-channel competition, Video-on-Demand

*I gratefully acknowledge the tremendous help I received from my supervisor Miguel Godinho de Matos. This work has been partially supported by the Portuguese Foundation for Science and Technology through research grant PTDC/EGE-OGE/27968/2017. We thank the industry partner for their support.

Analysis of sequential distribution on Video-on-Demand platforms

Laetitia Römer *

February 21, 2021

Abstract

A distribuição sequencial de títulos através de diferentes canais é uma estratégia prevalecente de maximização do lucro na indústria cinematográfica. No entanto, podem ocorrer efeitos entre canais, em que a distribuição de um produto num canal afecta as vendas num outro canal. Eu analiso como a disponibilidade de uma versão de um filme num canal digital impacta as vendas noutro canal digital no contexto de uma plataforma de Video-on-Demand que oferece títulos para alugar e para comprar. Utilizando dados secundários de uma plataforma de Vídeo-on-Demand, o meu estudo explora a entrada e saída sequencial das versões de um filme. Os meus resultados demonstram que para filmes disponíveis por mais de um ano antes de serem oferecidos como duas versões, após adicionar a segunda versão as vendas da versão pré-existente aumentam. Por outro lado, se os filmes estiverem disponíveis há menos de um ano antes de serem distribuídos como duas versões, as vendas da versão pré-existente diminuem. Concluindo, os resultados sugerem que dependendo do atraso entre a introdução da versão a alugar e a compra, a adição da versão a comprar ou tem um efeito de spillover positivo ou um efeito de canibalização.

Keywords: Versioning, Digital channel, Cross-channel competition, Video-on-Demand

*I gratefully acknowledge the tremendous help I received from my supervisor Miguel Godinho de Matos. This work has been partially supported by the Portuguese Foundation for Science and Technology through research grant PTDC/EGE-OGE/27968/2017. We thank the industry partner for their support.

Contents

List of Figures

List of Tables

1 Introduction

WarnerMedia, one of the world's largest entertainment conglomerates, announced in early December 2020 that it would simultaneously offer its movies in theatres and on its own streaming platform, breaking with the strategy of so-called windowing (Swisher, 2020). A movie is commonly distributed first in theatres and only subsequently through outlets such as DVD, television, or Video-on-Demand (Kumar et al., 2014). Windowing thus is the movie industry's implementation of the concept of versioning, where firms offer versions of the same core product at different prices to different consumer segments to maximise profits (Varian, 1995). How to increase profits with movies has received increasing attention lately not only from the industry but also from scholars investigating different distribution methods. An essential consideration with respect to windowing is the timing of releases. In this context, scholars examine how the distribution of a movie in one channel affects sales in another channel. Prior studies have focused mainly on physical versus digital channels (see Danaher et al., 2010; Kumar et al., 2014; Smith and Telang, 2009) but rarely investigated this phenomenon across different versions in digital channels. Gong et al. (2015) is a rare example thereof, using an experiment to analyse the effect of sequentially introducing a digital rental and version to buy of a movie. I address this gap in the literature by researching how digital movie sales impact digital movie leases, relying on data from the staggered introduction of different versions of the same title on a Video-on-Demand platform instead of an experiment.

I analyse the effects on sales of the sequential distribution of a movie to lease and to sell on the same digital platform. I use secondary data from a Video-on-Demand provider offering titles to lease and to buy. I use the sequential entry and exit of movie titles to this platform, where titles are subsequently available as versions to lease and buy. I limit my analysis on the effects of adding a version to buy to an existing leasing version of a title. Estimating a negative binomial dummy variable regression model, I show that there seems to be either a positive information spillover or a cannibalisation effect depending on the age of movies in the catalogue upon adding a title's version to buy to the pre-existing leasing version. If movies are in the catalogue longer before being available as two versions, adding the version to buy increases unit sales of the pre-existing version to lease. I explain this by an increase in awareness of platform users for the title from adding the second version. Previous research accordingly suggests that the promotion of one title in a channel leads to more awareness and hence increases sales in another channel (Gong et al., 2015). Movies added more recently to the platform do not seem to benefit from the increase in consumer awareness. Instead, possibly, a cannibalisation effect for the pre-existing version predominates: For movies distributed as two versions after a shorter time of availability in the catalogue, sales of the pre-existing version decrease upon the introduction of the second version. I use propensity score matching to show robustness of these findings indicating that, on average, subsequently adding a movie's version to buy increases sales for the pre-existing leasing version.

This thesis makes several key contributions. It is one of the few works to investigate the sales performance of the sequential distribution of different versions of a digital product in the movie industry. This thesis adds to a growing academic literature analysing the interaction between digital distribution channels. I extend this literature by using sales data collected from a provider in the Video-on-Demand industry. The empirical results offer managerial implications for platform providers and potentially firms in other creative industries. The critical implication is that sequential distribution is a viable strategy for increasing awareness among customers for a title. However, there is a potential cannibalisation effect from adding a different version on the platform as well. Therefore, platform managers should consider promoting and raising awareness for more recently added titles without necessarily adding them as an additional version to forego cannibalisation.

This thesis is organised as follows: section 2 reviews the literature on versioning and cross-channel cannibalisation. Section 3 describes the data, the model specification and empirical methods. I report results of the regression in section 4 and robustness checks in section 5. I discuss results in section 6, point out limitations in section 7 and conclude in section 8.

2 Literature Review

Ulin (2013, p. 3) notes that the "exploitation of media is a symbiotic process, where success is achieved by choreographing distribution across time and distribution outlets to maximise an ultimate bottom line". Versioning, which Ulin describes as the choreography necessary, and the resulting channel competition, are at the core of the movie industry's management challenges and, consequently, its distribution platforms, such as Video-on-Demand platforms.

Versioning is a business strategy used to maximise profits. If a firm's consumers differ in their willingness-to-pay for a product, it is preferential for the firm to price discriminate (Varian, 1995). Offering the same product at different prices allows a firm to maximise the profits it gains from each customer. However, for this strategy to work, a firm must differentiate between its customers to offer the rightly priced version in line with the customer's willingness-to-pay. It would be optimal for a firm to offer a different price according to each customer's willingness-to-pay, requiring knowledge of the latter. One possibility for a firm lacking this information is to segment customers by inducing them to self-select (Varian, 1997). A firm offers the same core product at different prices, with customers themselves choosing the version according to their marginal willingness-to-pay. By applying this second-degree price discrimination, the firm is, however, risking cannibalisation (Shapiro and Varian, 1998). Customers might defer to buying the lower-priced version instead of the higher-priced one with the seller consequently losing out on profits. To avoid cannibalisation of the highest-priced version, a seller can offer the product at different levels of quality. Such differences in quality can stem, among others, from varying delay (e.g. offering a hardcover version of a book before its paperback version or a movie first in theatres and then on DVD), convenience (e.g. restricting time, place or length where the

consumer can access the Internet) or resolution (e.g. offering stock photos at a higher or lower resolution) (Shapiro and Varian, 1998). In that way, a firm can keep customers from purchasing a version at a price that is not intended for them: A customer with a higher willingness-to-pay will choose the higher-quality product and pay a higher price. The different versions thus respond to different customer needs, allowing the seller to offer the same core information at different prices (Shapiro and Varian, 1998; Varian, 1997).

Within the movie industry, versioning is achieved by so-called windowing, which refers to the sequential distribution of a movie through different outlets. Windowing is a strategy used within the movie industry to maximise profits in the distribution of movies (Smith and Telang, 2016a) and affects Video-on-Demand platforms as to when titles are distributed. Different windows address different consumer segments. Typically, a principle of "second-best alternative" (Vogel, 2020, p. 147) is applied: A movie is first distributed where it attracts customers with the highest willingness-to-pay and the most marginal revenue, before showing it in lower-value channels to earn revenue from customers with a lower willingness-to-pay. Therefore, a movie is released in theatres first, followed by releases on DVD, Video-on-Demand and digital channels, before its distribution on pay-television, subscription pay-cable channels and finally on free-television (Kumar et al., 2014; Smith and Telang, 2016b). These established windows are increasingly under pressure from digitisation and globalisation. Consumers around the world are now more aware of movies that are first distributed in US theatres or online channels but do not have access to the same movies. Digital piracy, the illegal sharing of movies online, is offering consumers a way out, but at a high cost for the studios. It is a crucial reason why studios start to modify release strategies[1] (Smith and Telang, 2016b; Ulin, 2013).

Within (and not only across) windows and channels, quality, usability and price vary to increase product segmentation further. For example, studios offer movies in high definition (HD) or standard definition (SD) or as a version with or without bonus content (Smith and Telang, 2016b). The Video-on-Demand platform analysed within this thesis uses such differences, further segmenting its products besides a temporary delay of different versions.

An important management decision within versioning revolves around the optimal timing of the different version's distribution. While studios must time theatrical, DVD and TV releases, Video-on-Demand platform providers must decide whether to introduce different versions of the same title simultaneously or sequentially and if so with which delay. Since this thesis investigates the effects of sequential distribution, I highlight relevant literature on simultaneous and sequential distribution in the following.

The early work of Moorthy and Png (1992) discusses whether a seller should introduce two versions of a product (one of lower and one of higher quality) simultaneously or sequentially. The study's analysis is based on a model of a single-channel market with durable goods and

[1] WarnerMedia's decision to distribute versions simultaneously, instead of sequentially, is said to be partly driven by the Covid-19 pandemic with decreasing revenue from theatres, but mainly from trying to shift business towards online channels (Swisher, 2020).

increasing marginal costs of production. By offering versions simultaneously, the lower-quality version can cannibalise the higher-quality version. They show that delaying the low-end version prevents this potential cannibalisation. Further, they find that sequencing is more profitable than simultaneous release if the consumers are more impatient than the seller. However, since they assume increasing marginal costs of production, this model does not necessarily fit information goods and as such movies, which are defined by high fixed costs and low, if not zero, marginal costs (Bhargava and Choudhary, 2008). Within more recent literature and the movie industry, it is still controversial whether the sequential distribution of movies is more profitable than the simultaneous distribution [2] (Eliashberg et al., 2006). Arguments for and against sequential distribution revolve around whether the different distribution outlets are substitutes or complements. In case consumer segments are overlapping, the outlets are substitutes. Then, sales made through a channel that distributes movies early, possibly cannibalise sales in a later channel. If consumer segments are sufficiently distinct, the outlets are complements. Then, rather than cannibalisation, a spillover effect prevails. Sales in a later channel increase from marketing efforts or word-of-mouth revolving around the release of the movie in an earlier channel (Weinberg, 2005).

Choosing the sequential distribution of content[3] entails deciding on the optimal delay between versions. The earlier works of Frank (1994) and Lehmann and Weinberg (2000) focus on how to time the sequential distribution of theatre and video rentals optimally. Frank (1994) finds that on the one hand, releasing a video version of the movie too early leads to consumers not going to theatres since they can watch it at home. On the other hand, releasing it too late decreases demand from consumers since the movie is too old (measured in terms of theatre premiere). In line with this, Lehmann and Weinberg (2000) show that while the likelihood of cannibalisation increases with decreasing delay between the release in the two windows studied, extending the delay leads to a decrease of the impact the theatrical releases' marketing campaigns have on the second window, which diminishes sales. Within this thesis, the Video-on-Demand platform analysed uses different delays for introducing versions. Although in a different context (theatrical versus home video sale rather than digital leasing versus sale) previous literature shows that the delay between two versions' availability seems to affect sales and potentially increase (or decrease) cannibalisation. Therefore, in the following, I study potential heterogeneous effects from a different delay between offering a title's versions on the platform.

More recent studies question whether the sequential release of movies is profit-maximising for studios. Hennig-Thurau et al. (2007) and Calzada and Valletti (2012) show that simultaneous releases in theatres and on home video can be optimal for studios, with the side-effect of hurting theatres (Hennig-Thurau et al., 2007) or under the constraint that studios are integrated with the

[2] Within the industry, the most recent example for this discussion is the earlier mentioned announcement of WarnerMedia to offer movies simultaneously online and in theatres as of 2021.

[3] like the Video-on-Demand platform providing the data for this analysis does for at least some of the movies distributed

distribution channels[4] (Calzada and Valletti, 2012). Mukherjee and Kadiyali (2011), testing cross-channel substitution between DVD purchases and rentals, show that revenue decreases under sequential distribution. They find that when faced with older movies in their preferred channel compared to the other channel, consumers do not switch channels but rather defer from buying or renting from their preferred channel, respectively.

Besides versioning and the accompanying decisions on whether to sequentially distribute versions and if so with which delay, channel competition is another main management challenge within the movie industry and for Video-on-Demand platforms. An essential consideration, therefore, is how the distribution in one channel affects sales in another channel. Scholars predominantly examine the context of physical versus digital channels. Smith and Telang (2009) and Kumar et al. (2014) focus on the interaction between pay-cable broadcasts and DVD sales, surprisingly showing that DVD sales for broadcasted movies increase, rather than result in their cannibalisation. This positive spillover effect from an increased consumer awareness is larger for less popular or lesser-known studios (Kumar et al., 2014). Based on this finding, I investigate whether the sequential distribution of versions has a different effect on movies with higher IMDb[5] rating, which I choose as a proxy for a title's popularity[6]. Danaher et al. (2010) also study the cross-channel relationship between digital and physical sales, using iTunes and DVD box sales. They show that digital channels might cannibalise physical sales, but that there might also be enough differentiation across channels so that the goods respond to different consumer needs. Knox and Eliashberg (2009) model whether consumers rent or buy at a store offering DVDs for both rent and sale and develop pricing strategies for such a "rentailer" (p.125). Contrarily to the previous studies, they assume that upon entering the store, the consumer already knows which title he wants to consume and only decides whether to rent or to buy. This setup, however, means that only a cannibalisation effect can occur, since promotion in one channel could only lead to a negative impact on the other channel. Gong et al. (2015) is one of the few works studying the cannibalisation across digital channels. They use price discounts to show cross-channel effects between digital purchase and rental markets. Consumers are assumed only to purchase the same movie title once, making the different channels substitutes at the movie level and expecting cannibalisation between channels. However, they find that promotion in the purchasing channel using a price discount leads to an increase, rather than a decrease, in sales in the rental channel for the promoted movie. A potential information spillover between channels exists, where the promotion of a title in one channel raises the awareness for it (by receiving attention on blogs, online discussions or third-party websites), increasing sales in both channels. Similar to Gong et al. (2015), I analyse the effect of sequentially introducing a digital leasing

[4]An integrated studio not only produces but also distributes the movie (instead of using independent distributors).

[5]The Internet Movie Database (IMDb) provides information such as release date, genre and user ratings (on a scale from one to ten) on, among others, movies and television.

[6]This is in line with, for instance, Chintagunta et al. (2010) who show that the valence of online user ratings positively impacts box office earnings, indicating that the movie attracts a large number of viewers.

version and a version to buy. However, instead of using an experiment, I rely on the staggered introduction of different versions of the same title on the platform. While Kumar et al. (2014) or Danaher et al. (2010) focus on digital versus physical channels, I investigate whether spillover effects also persist across different digital sales channels.

3 Data and Empirical Methods

3.1 Data

I use secondary data to perform the analysis. The data is provided by a large multinational telecommunications provider, which I call TELCO in the following, for a period between January 2012 and February 2019. TELCO serves more than one million households in the analysed country. Besides Pay-TV, internet and telephony, TELCO offers a Video-on-Demand system. For several years, TELCO has hosted both old and new content on its platform and has offered content to lease (as temporal Video-on-Demand (tvod)) and to buy (as perpetual Video-on-Demand (pvod)). The different types of introduction of movies are visualised in Figure 1. TELCO releases titles only to lease, only to buy or to lease and buy on its platform. If the provider introduces a title as only either to lease or to buy, they might decide to add the other version with different delay. Similarly, TELCO sometimes releases both versions of a title simultaneously but removes one version before it removes the other. Furthermore, there are titles on the platform which are available as both versions but never simultaneously. I limit my analysis to movies introduced as tvod first and subsequently as pvod, which is the type of introduction highlighted in Figure 1. This limitation is due to both fewer movies available for the other types of introduction and my time constraint to complete this thesis. I do not have information on TELCO's decision-making process on why or when to add or remove the second version of a title to the platform.

Titles used within the analysis[7] have one spell[8] only, meaning that they enter the catalogue and either leave it at a specific date or enter the catalogue and remain in the catalogue indefinitely. A movie that is eventually available as a second version only remains within the dataset if it is available as one version first, before being added as a second version, meaning that movies are removed from the dataset that are made available as two versions directly upon entry in the catalogue.

I have daily data per movie title on the quantity sold and, if the title sold at least once on a specific day, on the price. Hence, if a movie did not sell on a respective day, I have no information on its offered price. I then use catalogue data, indicating the period within which a

[7] Titles haven been removed from the dataset for the analysis due to their movie-length being indicated at zero, not selling at least once, not having an unambiguous start and end date or being a duplicate entry.

[8] A spell is a distinct period of time a movie is available on the platform. E.g. a movie with two spells enters the catalogue, remains for a certain period, leaves the catalogue and after a certain time enters the catalogue again before leaving it again.

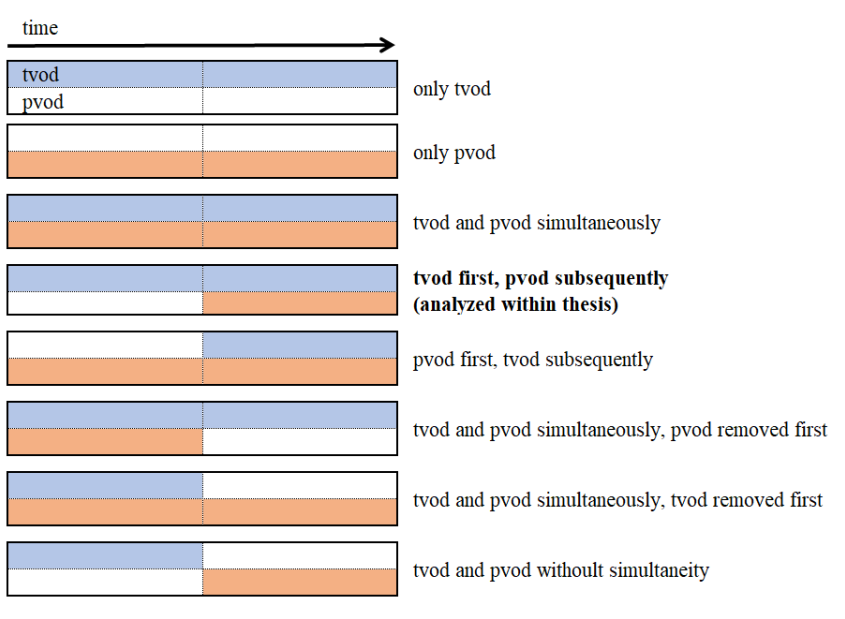

Figure 1: Types of introduction of movies on Video-on-Demand platform

title is available on the platform, to infer that if the movie is available but not recorded within the sales data on a specific date, the quantity sold on that date is zero. I aggregate daily sales data at a monthly level per title and version to reduce noise in the analysis. If a movie never sold during a month but was available on the platform, I fill the missing price with the last month's price, if available. Additionally to the aforementioned data, a dummy variable indicates whether a title is ever available as both tvod and pvod and another dummy variable *two versions* indicates a title that is simultaneously available as both versions during a month. For each title, I have information on its duration, genre and whether it was available to display in HD.

For the period between February 2014 and August 2016, I have daily data on the location of a title's version on the platform[9]. The platform consists of different menus and sub-menus, and, within those, titles appear at a particular rank. To reach a specific sub-menu, a platform user must click a particular number of times, depending on the sub-menu's level. Inside a sub-menu, a platform user sees the first six titles and must click again to see the subsequent six titles at less visible ranks. A title can appear in different sub-menus simultaneously and change ranks within each sub-menu. I summarise the data at a monthly level to match it with the sales data. I use the minimum number of clicks necessary per month to find the title and retain only this value for the specific month. The number of clicks necessary consists of combining the sub-menu's level and the movie's rank within the sub-menu. An increasing amount of clicks necessary to find a title suggests that the movie is more difficult to find on the platform. I use the minimum clicks necessary across both tvod and pvod versions of a title since platform users can still decide which version to consume upon clicking on the title. I include the summarised data as a control variable for the title's location in my regression models.

[9]Data on the location of movies on TELCO's platform was only recorded for this limited amount of time.

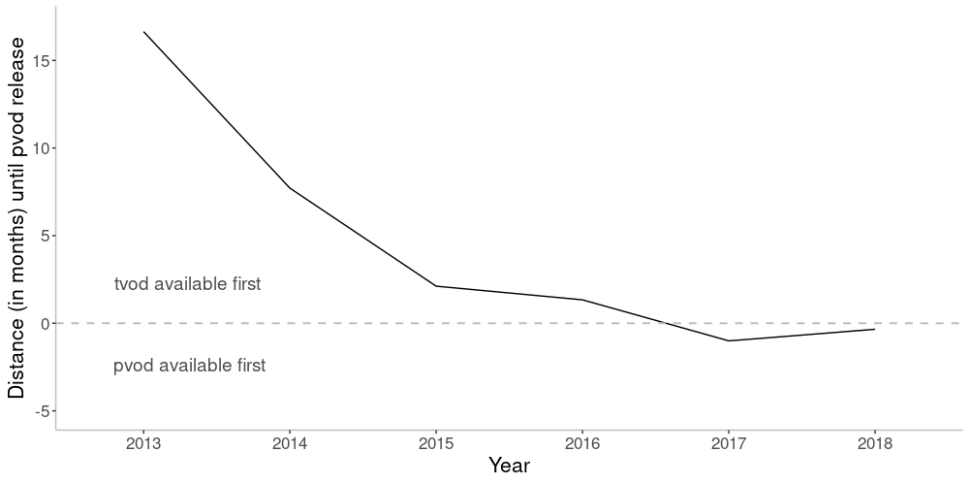

Figure 2: Lag between introducing versions for titles eventually available as both versions

3.2 Descriptive Statistics

I observe sales data between January 2012 and February 2019 in the final dataset. Until September 2013, only tvod titles are available on the platform. After the introduction of pvod versions, the proportion of clients buying pvod increases, but remains lower compared to sales made in tvod versions. In 2014, the first year pvod versions are available throughout the whole year, pvod sales account for 0.16% of sales of movies only available as one version, while until 2019 this proportion increases steadily to 14.2%. For movies available as both versions, in 2014 pvod sales account for 1.59% of sales of movies available as two versions, increasing to 51.1% in 2018 and 35% in 2019. Most titles on the platform which eventually become available as two versions are introduced as tvod first, but the proportion of movies introduced as pvod first increases over time. Figure 2 shows the mean time lag between introducing a title's tvod and its pvod version by year of introduction of the first version on the platform. The positive but decreasing distances in years 2013 to 2016 indicate that the tvod version of a title enters the platform first, while the delay between offering the versions on the platform decreases. The negative distances in years 2017 and 2018 indicate that the order of introducing a title's versions changes: on average, the pvod title enters the platform first, while the tvod version is delayed[10]. On average, tvod titles are available on the platform for 50 months, pvod titles for 31 months. This difference in the period under observation stems from the fact that pvod titles are only available on the platform from 2013 on. The delay between the introduction of two versions varies greatly: A tvod title is available as a second version after 23 months on average, with a standard deviation of 21 months. Therefore, and since previous literature uncovers different effects on sales and cannibalisation with the varying delay of versions (see section 2), in the model, I investigate whether heterogeneous effects based on how long a movie has been in the catalogue before being available as two versions occur (see section 4.2).

The total number of sales per year show a declining trend while total revenue shows no

[10]No movies available as both versions enter the catalogue in 2019; therefore the year is omitted in Figure 2.

Table 1: Average monthly unit sales and prices - all titles

	all	tvod only	tvod, also pvod	pvod only	pvod, also tvod
Titles					
number	3899	2821	338	575	165
Quantity Sold					
min	0	0	0	0	0
max	6708	6708	3441	653	873
mean	5.39	5.86	5.91	1.27	4.59
sd	48.85	53.29	42.91	11.22	26.85
Price					
min	0.49	0.49	0.99	2.99	5
max	20.99	5.29	4.49	20.99	19.99
mean	3.46	3.06	3.2	12.4	14.87
sd	2.32	1.12	0.74	3.75	2.97

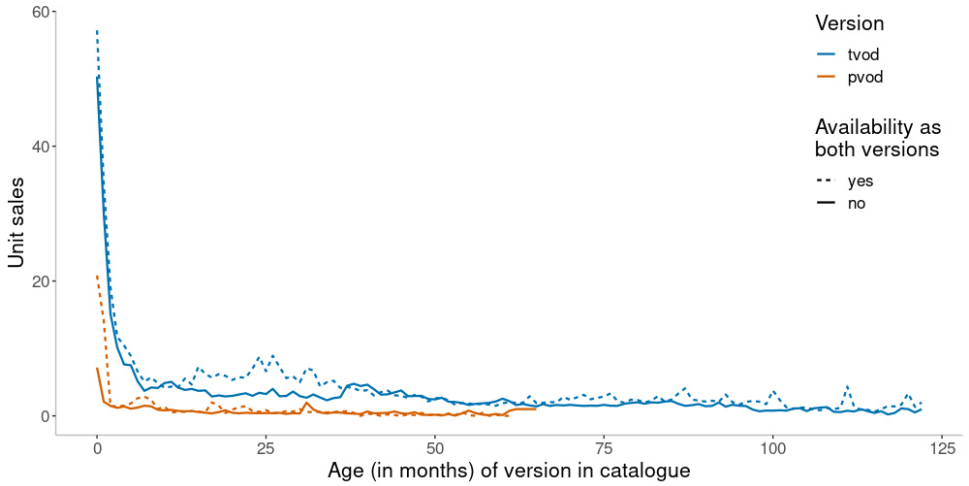

Figure 3: Average monthly sales trend for different types of availability of versions

clear trend, indicating increasing prices, which holds across both titles selling as one version and those selling as two versions. As observable in Table 1, on average, a movie sells 1.27 times within one month if it is pvod and available as one version only, 4.59 times if pvod and also available as tvod, 5.86 times within one month if it sells as tvod only and 5.91 times if the tvod title is also sold as pvod. The price per title is on average €12.4 if it is pvod and available as one version only, €14.87 if pvod and available as tvod, €3.06 if the title sells as tvod only and €3.2 if the tvod title also sells as pvod. Figure 3 depicts the average sales trend for the previously described types of availability by the number of months since their introduction to the platform. While Table 1 shows that tvod versions generally sell more than pvod versions, Figure 3 visualises that the same applies independently of the duration since a version's introduction to the platform. Further, both tvod and pvod monthly unit sales are higher if a title eventually becomes available as both versions, compared to if only one version is ever available. There is no data available on the average monthly unit sales of pvod titles on the platform for more than 65 months (where the pvod line ends), since pvod versions only exist since 2013.

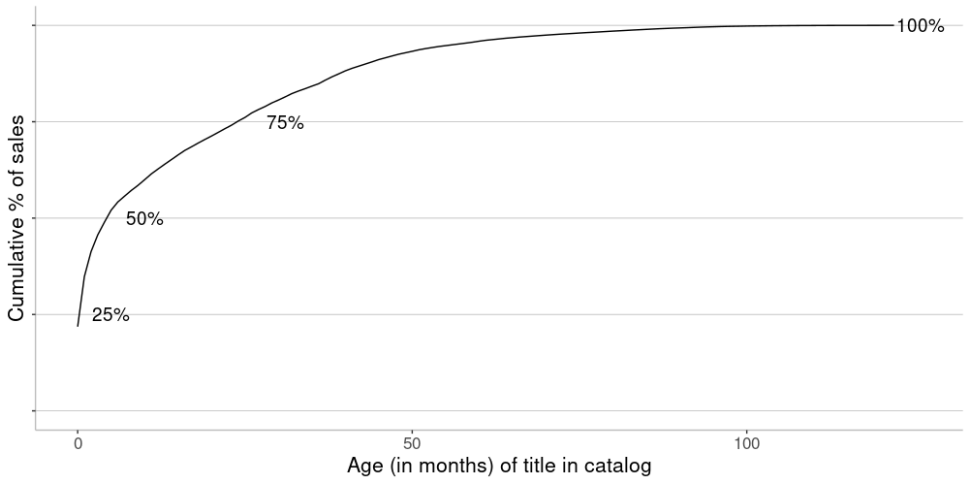

Figure 4: Cumulative sales with increasing age in catalogue

Monthly unit sales in the sample are highly skewed. Table 1 shows that monthly unit sales for tvod titles also available as pvod range from 0 to more than 3,400 with standard deviations seven times as large as the sample means. Similarly, Figure A1 in the Appendix visually shows the right-skewed distribution of tvod monthly unit sales. The majority of observations amount to zero or little more than zero monthly sales units while very few observations attain a large number of monthly sales. The skewed distribution of sales is typical in the movie industry: A small number of titles gain a disproportionately large proportion of, for example, theatrical or DVD sales, which Kumar et al. (2014) explain by consumers only being aware of a limited number of movies available.

On average, within one month of availability in the catalogue, titles sell 35% of their lifetime unit sales in the catalogue and 52% of sales within the first five months. Figure 4 depicts this decreasing increase of cumulative sales by the number of months a movie is available on the platform. The observation is in line with previous literature, showing that products sell less over time (Frank, 1994; Hennig-Thurau et al., 2007; Waldfogel, 2016). I, therefore, control for a title's age in the catalogue in the model, indicating the duration it has been available on the platform. Since I observe a decrease in sales with increasing age at a decreasing rate, pointing to a non-linear relationship between the regression model's outcome variable unit sales and the movie's age in the catalogue, I include the squared age in catalogue in my models.

The subset used in the following analysis consists of movies available as two versions simultaneously at some point during their lifetime within the catalogue. On average, titles within this subset sell less monthly units when available as two versions compared to when available as one version only. Table A1 in the Appendix shows that on average, in line with Table 1, tvod titles sell more units per month than pvod titles (however only slightly so if two versions of a title are available on the platform). Tvod titles eventually available as two versions sell, on average, nearly three times more when only available as tvod. Similarly, pvod titles eventually

available as two versions sell, on average, nearly twice as much when only available as pvod[11].

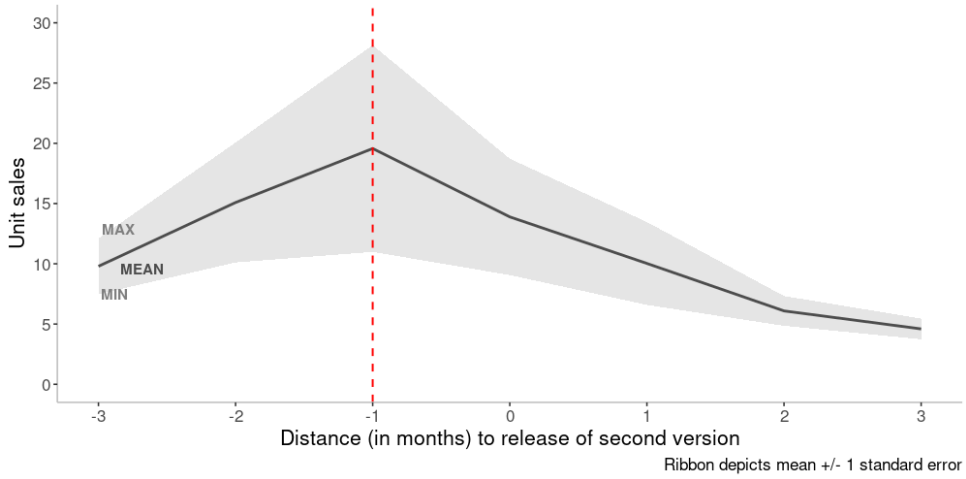

Figure 5: Tvod monthly sales trend

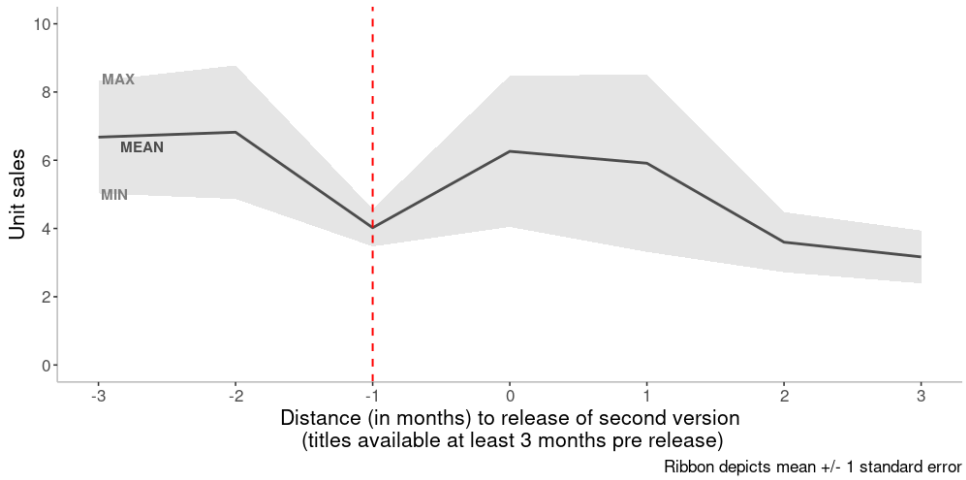

Figure 6: Tvod monthly sales trend for titles in catalogue 3 months before availability as second version

Figure 5 visually compares average monthly sales of tvod titles before and after adding a pvod version of the title to the platform on a relative time scale with zero indicating the moment the second version becomes available. The grey ribbon depicts the mean plus and minus one standard error, respectively. Adding a second version seems to decrease tvod sales: from relative month zero onwards, monthly unit sales, on average, are always smaller than the previous month. Surprisingly, assuming that with the increasing age of a title on the platform unit sales decrease, the average monthly unit sales before the availability of the second version are increasing. A possible explanation is that some titles become available as tvod only shortly before becoming available as two versions. There are thus titles within this subset which are

[11]However, this calculation of means does not account for the previously noted effect that older movies sell less, on average. Since tvod versions within this dataset are available as tvod first and only later as pvod (while, likewise, pvod versions are available as pvod first and only later as tvod), it is not surprising that the tvod and pvod versions sell more, respectively, if offered as sole version.

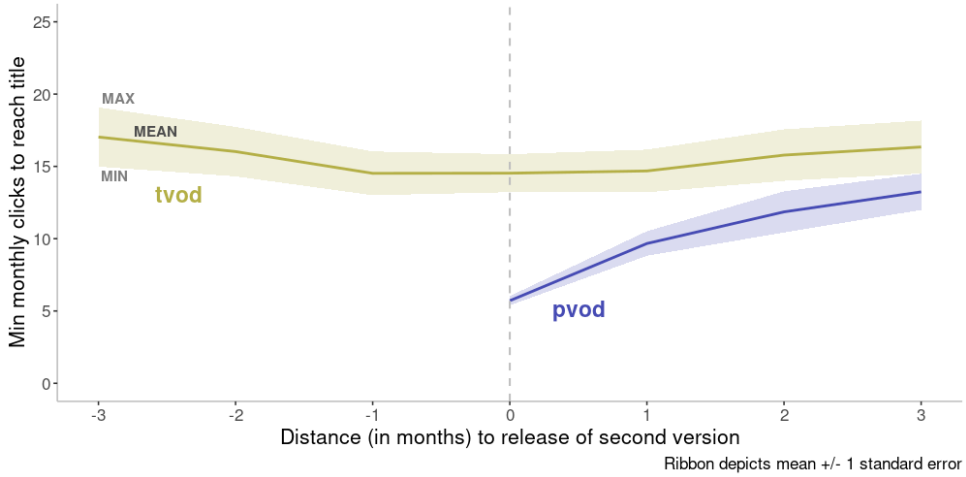

Figure 7: Rank within menu of tvod and respective pvod version

introduced to the catalogue and only one month later become available as pvod. Since recently added movies sell more, they possibly increase the average and obscure the overall negative sales trend with increasing age. To catch the effects of the second version's introduction, I, therefore, exclude such titles and use a subset of movies available for at least three months as tvod only before becoming available as a second version as displayed in Figure 6. In this Figure, I observe a general declining trend in sales over time with a peak once the second version is added at relative time zero. The generally decreasing sales are in line with the observation that movies sell a large proportion of total sales within a short period after being added to the catalogue and that movies sell less with increasing age.

Within the regression, the analysis is limited to this subset of sales data three months before and after becoming available as second version to catch the effects of the introduction of the second version. Since I am only considering movies that sold at least once during the chosen period, I calculate the regression on 294 tvod titles.

Using the location data available on a subset of movies, Figure 7 shows the minimum monthly clicks necessary to reach a title on the platform by relative time until the introduction of the second version for tvod and pvod versions, respectively. Zero indicates the introduction of a pvod version to a pre-existing tvod version. The figure shows that, on average, the tvod title's location remains stable across time. The pre-existing tvod version does not move to a more visible location upon introducing the pvod title. However, a title's pvod version enters at a more visible rank or in a more easily reachable sub-menu than the respective tvod version. This more visible location is only temporary: the minimum monthly clicks necessary to reach a title's pvod version converge towards the tvod version's location within a few months of the version's introduction. The data available also indicates that movies at a smaller, more visible rank sell more monthly units, on average, than movies that are more difficult to find. Therefore, I control for the title's location within the regression models.

3.3 Empirical Strategy

To investigate the research question of how the introduction of a pvod version impacts the sales of the pre-existing tvod version of a movie, I rely on the sequential entry of movies on the Video-on-Demand platform as versions to lease and buy. I measure the impact of a release of a title's second version on the monthly unit sales of the pre-existing version, focusing on the tvod version as pre-existing version. All titles used for the analysis are eventually available as two versions, meaning that I have a staggered implementation. I compare unit sales of tvod titles which are already available as both versions against sales of those titles still only available as tvod, but which will eventually become available as a pvod version, too. The titles in the sample are changing the number of versions available only once since I previously removed all titles that change the number of versions available more than once.

As shown in section 3.2, the distribution of the outcome variable monthly unit sales in the sample is highly skewed. A common approach to account for an outcome variable with a skewed distribution is to use a logarithmic transformation of the outcome variable within an ordinary least squares regression. However, since more than 31% of the data are zero-unit sales, I would have to drop a considerable amount of data. Furthermore, the outcome variable monthly unit sales is count data, taking on non-negative and integer values. The conditional mean and the conditional variance of monthly unit sales are different within the data: conditional on a title available as one version or two versions, mean monthly unit sales amount to 15.27 and 8.82, respectively while standard deviations are 103.29 and 53.57, respectively[12]. Therefore, I follow another commonly used approach (for several references see Gong et al., 2015), where a negative binomial regression predicts the expected log count of the outcome variable by using a linear combination of predictors. The model equation takes the following form:

$$log(\widehat{Q_{it}}) = b_1(two\_versions_{it} = 1) + b_2 age\_catalogue_{it} + b_3 age\_catalogue_{it}^2 + b_4 price_{it} + b_5 location_{it} + \alpha_i + \theta_t + \varepsilon_{it} \quad (1)$$

where $log(\widehat{Q_{it}})$ is the expected log count of the number of sales of movie i during month t. The dummy variable $b_1(two\_versions_{it} = 1)$ is equal to 1 if the movie is available in two versions and is the coefficient of interest. It represents changes in sales of a movie i when the second version is added to the catalogue. If positive, the coefficient of interest indicates that upon introducing the pvod version of the movie, the tvod version's sales increase on average. If negative, the title's tvod version sales decrease on average when additionally offering the pvod version on the platform. The variable $b_2 age\_catalogue_{it}$ indicates the number of months from the release on the platform of movie i until the month t to control for a movie's age. Since Figure 4 depicts a decrease in sales with increasing age of a movie on the platform but at a decreasing rate,

[12] If conditional mean and standard deviation were equal, I could use a Poisson regression to model the count variable. Since this does not apply for my data, I use a negative binomial model regression, which is similar to a Poisson regression except that it uses an additional parameter to model the over-dispersion.

pointing to a non-linear relationship, I also include the squared age in Equation 1. To control for changes in price, I include the variable $b_4 price_{it}$, indicating the price of a movie i during month t, since theory predicts that an increase in price leads to a decrease in demand (Blaug, 1997). Controlling for price is especially relevant, given that digital movies' own-price elasticity is higher than price elasticities for physical products (Gong et al., 2015). Therefore, prices are likely to influence consumers on Video-on-Demand platforms in their decision-making. Figure 7 depicts that the minimum clicks necessary to find a movie decrease considerably upon the introduction of the second version. On average, the movie's second version enters the platform at a more visible location compared to it's pre-existing version. Therefore, I include the variable $b_5 location_{it}$ in my model to control for the location of movie i during month t on the platform. α_i represents movie fixed effects, and θ_t represents time (year and month) fixed effects.

I interpret the parameters estimated by maximum-likelihood-estimation in the negative binomial regression like a linear regression model's coefficients, where the outcome variable is in a $log(y)$ form. Since I use panel data, the error term ε composites time-invariant errors u_i and errors varying by time ε_{it}. Using a dummy variable regression approach by including a dummy for each year, month, and title allows eliminating all time-invariant errors u_i. This approach is mitigating the risk of endogeneity due to unobserved time-invariant variables. The remaining errors varying by time ε_{it} may still be heteroskedastic, autocorrelated or serially correlated. Therefore, I specify cluster robust standard errors on movie titles to handle those issues and to allow for valid inferences (Cameron and Miller, 2015).

To investigate whether certain movies' characteristics moderate the effects of adding a second version of a title to the catalogue, I extend the model by including interaction terms between the coefficient of interest and different movie characteristics (see section 4.2). I test whether the delay between a movie first entering the catalogue and becoming available as two versions or its IMDb rating change the effects of becoming available as two versions, respectively.

4 Results

4.1 Main Regression

I report the basic regression model results regarding the impact of adding a pvod version to a tvod title in Table 2. For better traceability of my results, I include six columns, each with the same dependent variable monthly unit sales and always including movie ID, year and month dummies. Column (1) serves as a reference, showing only the coefficient of interest $b_1(two\_versions_{it} = 1)$. Column (2) additionally includes a control for the movie's price and column (3) adds a control for age in the catalogue to the model. In columns (4) to (6) I calculate the model on a subset of data for which information on a title's location is available. Since this information is only available for a limited amount of time, columns (4) to (6) show results based on fewer titles and thus fewer observations than the previous columns. Therefore, column (4)

serves as a reference for the following models based on this subset. Column (5) controls for the location of the movie in the menu while column (6) finally combines the different controls. Due to the log-transformation of the outcome variable monthly unit sales, I interpret the effect of $b_1(two\_versions_{it} = 1)$ as the expected change in the log of monthly unit sales for a one-unit increase in *two versions* holding all other variables fixed. Hence, upon changing *two versions* by 1, I expect the outcome variable to change by $\%\Delta y = 100 * (e^{b_1} - 1)$.

Table 2: Main regression results

	Dependent variable:					
	monthly quantity					
	(1)	(2)	(3)	(4)	(5)	(6)
two versions	0.052	0.028	0.037	0.099	−0.218	−0.212
	(0.117)	(0.112)	(0.114)	(0.242)	(0.229)	(0.224)
price		−0.956***	−0.754**			−1.724***
		(0.284)	(0.328)			(0.425)
age catalogue			−0.201***			−0.342***
			(0.035)			(0.115)
age catalogue^2			0.001***			0.002
			(0.0003)			(0.001)
location					−0.032***	−0.029***
					(0.009)	(0.009)
Movie ID Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,908	1,908	1,908	606	606	606
Log Likelihood	−3,874.316	−3,867.963	−3,846.241	−1,173.759	−1,158.641	−1,154.213
θ	2.487*** (0.146)	2.511*** (0.148)	2.642*** (0.158)	2.204*** (0.221)	2.513*** (0.262)	2.612*** (0.277)
Akaike Inf. Crit.	8,370.632	8,359.926	8,318.483	2,569.518	2,541.281	2,536.425

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Standard errors clustered by movie ID
Including dummy variable per Movie ID, Year and Month

The coefficient of interest in column (1) is positive but not statistically significantly different from zero, suggesting that if a movie is available as two versions, it does not sell more on average than if it is available as only one version. In column (3), the coefficient of interest remains not statistically significantly different from zero. Conforming to the law of demand (Blaug, 1997), the coefficient for price is negative and indicates a statistically significant decrease in sales with an increase in price. In line with what theory predicts, the coefficients for age indicate that with increasing age, a movie sells significantly less, while this decrease in unit sales occurs at a decreasing rate. Column (5) shows that with an increase in clicks necessary to find a movie within the Video-on-Demand platform menu (hence with a less visible location), a movie sells significantly less monthly units, on average. While the coefficient of interest is now negative, indicating that it is affected by location since the coefficient of interest is positive in the previous column excluding controls, it remains not statistically significantly different from zero. Column (6) shows the results based on a model including controls for price, age and location in the menu. Adding a pvod version, on average, still has no statistically significant difference on monthly tvod unit sales. The coefficients for price, age and location remain unchanged in direction. I cannot infer from those results that adding a second version of a movie title changes the title's tvod sales on average.

4.2 Heterogeneous Effects

I further examine whether the impact of adding a second version to the platform varies across different characteristics of the titles. Previous literature indicates that the delay between two versions' availability seems to affect sales and potentially increase (or decrease) cannibalisation (Frank, 1994; Lehmann and Weinberg, 2000). Therefore, I test whether the delay between a title being added to the platform as tvod version and becoming available as a second version moderates the effect of a title's availability as two versions. To conduct this analysis, I add interaction terms between the second-version dummy *two versions* and the dummy variable *delay* $\geq$ *4 months* as well as the dummy variable *delay* $\geq$ *13 months* to the list of explanatory variables. The additional dummy variable *delay* $\geq$ *4 months* indicates a title being available between four and twelve months on the platform before becoming available as a second version. The dummy variable *delay* $\geq$ *13 months* indicates titles that are available for thirteen or more months in the catalogue before becoming available as a second version. I choose those periods for cut-off according to the distribution of sales across the movies' lifetime sales on the platform, as shown in Figure 4. The model is estimated with the same dependent variable as before, log-transformed monthly unit sales. I report the extended model's result in Table 3. Similar to the basic regression model results in Table 2, I report results for different control variables and combinations of control variables to enhance the results' traceability. The control variables' directions included throughout columns (2) to (6) remain unchanged compared to the basic regression model results.

Table 3: Availability as two versions interacted with delay until second version

	Dependent variable:					
	monthly quantity					
	(1)	(2)	(3)	(4)	(5)	(6)
two versions	-0.727^{***}	-0.733^{***}	-0.661^{***}	-0.699^{**}	-0.783^{***}	-0.809^{***}
	(0.159)	(0.157)	(0.160)	(0.290)	(0.281)	(0.284)
price		-0.693^{**}	-0.680^{**}			-1.385^{***}
		(0.307)	(0.316)			(0.391)
age catalogue			-0.159^{***}			-0.110
			(0.033)			(0.071)
age catalogue^2			0.001			-0.00005
			(0.0004)			(0.001)
location					-0.025^{***}	-0.025^{***}
					(0.009)	(0.009)
two versions x availability $>$ 4 months	0.643^{***}	0.666^{***}	0.667^{***}	0.374	0.278	0.282
	(0.169)	(0.169)	(0.169)	(0.370)	(0.374)	(0.376)
two versions x availability $>$ 13 months	1.357^{***}	1.324^{***}	1.163^{***}	1.351^{***}	1.089^{***}	1.105^{***}
	(0.192)	(0.193)	(0.212)	(0.376)	(0.392)	(0.385)
Movie ID Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,908	1,908	1,908	606	606	606
Log Likelihood	$-3,821.877$	$-3,818.365$	$-3,816.584$	$-1,157.242$	$-1,147.855$	$-1,145.529$
θ	2.822^{***} (0.173)	2.836^{***} (0.174)	2.842^{***} (0.174)	2.552^{***} (0.269)	2.754^{***} (0.297)	2.819^{***} (0.306)
Akaike Inf. Crit.	8,269.754	8,264.730	8,263.168	2,540.484	2,523.710	2,523.058

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Standard errors clustered by movie ID
Including dummy variable per Movie ID, Year and Month

The coefficient of interest in column (3), including controls for a title's price and age in the catalogue and the interaction terms, shows that, on average, adding a pvod version to a tvod title

decreases monthly unit sales by 48.37%, statistically significant at the 1% level, for titles with a short delay between the versions' availability. In contrast, the coefficient of the interaction term between a title available as two versions and a title being in the catalogue for four to twelve months before availability as two versions is positive and turns the effect of *two versions* to being positive. Therefore, on average, monthly unit sales for movies available as two versions after a medium period on the platform increase by 0.60%, statistically significant at the 1% level. Finally, the coefficient of the interaction term between a title available as two versions and a title being in the catalogue for thirteen and more months before becoming available as pvod is positive and larger than the previous interaction term's coefficient. On average, monthly unit sales for movies available for a long time before becoming available as a second version increase by 65.2%, statistically significant at the 1% level. The results suggest that adding a second version has a negative effect on movies that are available as a second version only shortly after their introduction to the platform. Contrarily, sales increase, on average, if the delay between the first and second version is sufficiently large.

When additionally controlling for location in column (6), results are mostly similar. The coefficient of interest remains negative and statistically significantly different from zero, indicating that sales of the pre-existing version decrease upon the introduction of the second version by 55.47%, on average. The interaction term's coefficient between *two versions* and a long delay until the second version's availability remains positive and statistically significantly different from zero, turning the effect of adding a second version positive as previously. The results indicate that for movies with a long delay between versions, monthly unit sales increase by 34.44% upon the introduction of the second version, on average. The coefficient of the interaction term between *two versions* and a medium delay is not statistically significantly different from zero. Thus, I cannot infer that, when additionally controlling for a title's location, a movie sells more or less monthly units on average if the second version's delay is between four and 13 months compared to when the delay amounts to up to three months. The location control and price coefficients remain unchanged in direction and magnitude. The controls for age, however, are now not statistically significantly different from zero, indicating, against previous assumptions, that age does not affect monthly unit sales if also controlling for a movie's location.

I further examine whether adding a second version of a title has a different impact on different types of movies. Since I can plausibly assume that the IMDb rating is not affected by adding a second version to the platform, I test whether this characteristic moderates the effects by adding the interaction term between the second-version dummy and IMDb rating. I re-estimate the model with the same dependent variable as before, the log of monthly tvod unit sales. I include the interaction term between the dummy variable *two versions* and IMDb in the model, reporting the results in Table A2 in the Appendix. The number of observations is smaller than previously since the IMDb rating is only available for a limited number of movie titles[13].

[13]While essentially any movie is available on the IMDb, titles within this dataset are not sufficiently distinguish-

If including controls separately in columns (2) to (6), the coefficients align with the respective columns in the basic regression model. If only including age and price as control, the coefficient of interest is positive and statistically significant at the 5% level, while the coefficient estimate on the interaction term is negative and statistically significant at the 10% level. The implication is that the effect of an additional version differs for different levels of IMDb rating. Sales of movies with a higher IMDb rating seem to benefit less from adding the second version since tvod unit sales increase less after a title becoming available as two versions. However, if adding the location control to the list of explanatory variables, neither the coefficient of interest nor the interaction term's coefficient is statistically significantly different from zero. Thus, I cannot conclude that the effect of an additional version differs for different levels of IMDb rating if controlling for the location of a movie in the Video-on-Demand platform's menu.

5 Robustness Check

I do not have information on the decision of TELCO to add the second version of a movie. While I use a dummy variable regression for my analysis, which allows me to eliminate time-invariant errors from the model, I do not control for time-variant unobservable factors. Optimally, to study the effect of adding a second version of a movie to a pre-existing title on the platform, I would conduct a randomised experiment by randomly picking movies to become available as a second version. Thereby, no unobserved variables would be influencing the decision of whether to add a second version of a movie or not. However, conducting such an experiment is not within the scope of this thesis. Therefore, I re-estimate the regression with a propensity score matching (PSM) method (for a similar approach see Smith and Telang, 2009) to additionally control for time-variant unobservable factors in the model. I use observable variables such as price and quantity sold during the first month of availability, genre, duration of availability in the catalogue and availability as an HD version to predict the probability of a movie's introduction as a second version to the platform. The matching allows comparing movies which have similar observable characteristics, meaning a similar propensity score, but where one movie is available as two versions on the platform, and one is not. The matching should also increase the similarity within unobservable factors between movies available as one and two versions, respectively, and thus help to control for time-variant unobservable variables.

I use a probit function to calculate propensity scores for the dependent variable (1 if a movie is ever available as two versions on the platform) with the observed explanatory variables. I subsequently use nearest neighbour matching as a method for matching movies based on their propensity scores. I use a subset of only those movies available during January 2018. I choose this subset of data since firstly, a relatively high proportion of movies are available

able from one another to extract the rating for each of them. For example, the IMDb has several entries for *The Raven*. There are two movies with the same title, which premiered in different years. Since I do not have data on when a movie was first published, I cannot unambiguously assign the relevant IMDb rating.

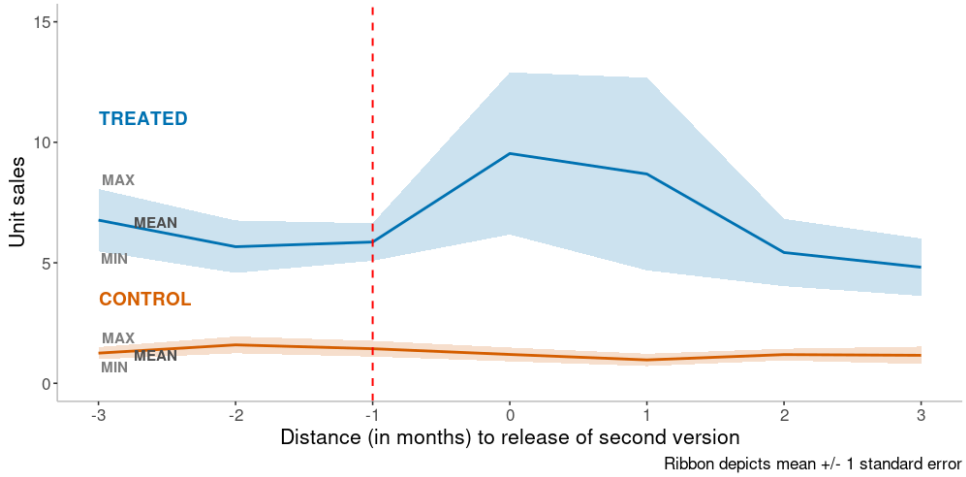

Figure 8: Tvod monthly sales trend - control and treatment group (PSM)

and secondly a relatively high proportion of movies become available as a second version in that month. To enhance similarity between movies, I use exact matching for the first month available. Therefore, I lose several movies available as two versions, since there are no movies available as only one version fitting those constraints. I further use a 1 to 1 PSM approach. 175 titles out of 191 movies from the group of movies distributed as two versions and available in January 2018 are matched. Since PSM relies on finding movies with similar propensity scores in both groups, I plot the covariate balance of the movies before and after matching in Figure A3 in the Appendix. I observe that the difference between the groups decreases after matching (indicated by points on the dots moving further towards zero upon matching) and that most differences after matching lie below the threshold of 0.1, indicating that the method allows us to compare more similar movies based on their observable characteristics.

I visually compare the sales of the control group (titles only ever available as one version) and the treatment group (titles eventually available as two versions) in Figure 8. For the control movies, I assign January 2018 as the month zero[14], while for treatment movies zero indicates the month in which the second version becomes available. I observe that movies in the control group sell less on average than movies in the treatment group and that pre-trend differences between the groups exist. While sales in the treatment group slightly decrease before the second version's availability, they remain stable within the control group. Upon the second version's availability, sales in the treatment group increase before declining again while the control group's sales remain constant, at least visually confirming the assumption that adding a second version to the platform impacts a title's sales.

I recalculate the regression analysis based on the movies from the newly defined treatment and control group. Differently to the regression results previously shown, I do not control for the location of the title on the platform within this analysis due to data constraints, as data on a

[14]since January 2018 is the month where the largest number of titles become available as a second version

title's location is not available for the relevant period[15]. The estimate on the difference in sales of movies before and after being available as two versions is positive and statistically significant at the 5% level as reported in column (1) in Table A3 of the Appendix. The direction of the coefficient remains unchanged compared to its calculation without PSM in the main regression if only controlling for price and age of the movie in the catalogue (see Table 2). Adding a pvod version to the platform leads, on average, to a 48.88% increase in sales of this title's tvod version. Most of the movies considered within the PSM have been in the catalogue for a long time before their second version's availability. Hence, the PSM regression results align with the previous finding that for movies with a long delay until the second version's release, the effect is positive.

To further increase similarity among the control and treatment group, I include the IMDb rating of movies as a covariate in the matching process. Since this information is only available for a limited number of movies in the subset, the treatment and control groups have 109 movies each after using exact and nearest neighbour method matching as before. I report the covariate balance before and after matching in Figure A4 in the Appendix. I observe that while the absolute standardised mean difference decreases upon matching as in the previous matching excluding the IMDb rating, the control and treatment movies are less similar since the number of possible matches is lower than before. Nevertheless, results of the visual comparison and the regression remain unchanged when adding the IMDb rating as covariate to the PSM. In line with Figure 8 depicting average sales trends for the PSM excluding IMDb rating as covariate, Figure A2 in the Appendix also shows that movies sell less on average if they are in the control group compared to if they are in the treatment group. Furthermore, there are still pre-trend differences between the two groups. Upon the second version's availability, sales in the treatment group increase before declining again while the control group's sales remain constant. Column (2) in Table A3 in the Appendix reports regression results. It indicates, in line with the previous regression results using PSM without the covariate IMDb rating, that if solely controlling for price and age, adding a pvod version to a tvod title leads to an increase in sales of the tvod version. Since most movies considered within the PSM have long been available before the treatment moment, the PSM regression results align with the previous findings.

6 Discussion

The results suggest that for movies with a long delay until the second version's availability, the second version's introduction has a promotional effect. Since consumers do not have perfect information, product discovery plays a role in their decision-making. Previous research shows that this is also the case in the movie industry, where the inflow of new products is high and therefore product discovery is a driver for consumer choices (Kumar et al., 2014). On online platforms, providers steer product discovery using recommendation systems to make consumers

[15] I further note that estimating the PSM for a period with location data available is restricted by data limitations.

aware of a tailored subset of the available products or place products more or less visibly on the platform. For Video-on-Demand platforms, this translates into placing a title closer to or further away from the front page, into menus which are more or less specific and easy to find and into placing a title at the top or bottom of the respective menu. This location of a movie within the platform changes the consumer's awareness of a title. In this context, Gong et al. (2015) show that promoting a title in one channel raises awareness for the title's "umbrella brand" (p. 344). Sales increase not only in the promoted channel but also in other channels where the title is available. Therefore, I attribute my results at least partially to the additional version raising awareness for the pre-existing version. However, data used within this analysis shows that pvod versions enter the platform at a more visible location than the respective tvod version is located simultaneously (see Figure 7). Independent of which title's version consumers encounter on the platform first, they choose the version they want to consume. Therefore, one version of a title becoming easier to find on the platform entails that the other version, as well, is more visible. To rule out the possibility that the increase in tvod sales solely stems from the title's more visible location within the platform upon the second version's availability, I introduce an additional control. When controlling for a title's location, the pvod version's availability still positively affects tvod sales for titles with a long delay between the introduction of the first and subsequent version. Sales seem to benefit from an increased awareness beyond the tile's location, such as a title's inclusion in a larger number of sub-menus once available as two versions.

For movies distributed as two versions with a short or moderate delay after first becoming available on the platform, there does not seem to be a similar benefit in sales. I attribute this lack of increase in sales to an already higher awareness for more recently added movies. Tvod versions introduced over one year ago do not benefit from the same awareness than titles added more recently, as the demand for movies generally declines with increasing time since its theatre premiere (Frank, 1994). Therefore, the second version's availability possibly changes the awareness more for an older than a newer title, explaining the difference in findings. Contrarily, the results suggest that adding a pvod version with only a short or moderate delay cannibalises tvod unit sales, where customers possibly defer to buying the pvod version instead. However, since I do not analyse the effects on pvod titles as the data available is too limited, this is only suggestive.

In line with the previous assumption that awareness for a title is a driver for the results, movies with a higher IMDb rating benefit less from the positive spillover effect from adding a second version to the platform. Assuming that more popular movies generally have a higher IMDb rating[16], such more popular movies could be benefitting from a generally higher awareness. Thus, adding a second version and the accompanying increase in attention would have less impact on those movies. This finding is also in line with a study conducted by Kumar et al. (2014), showing that a positive spillover effect in the context of DVD sales and broadcasted

[16]since previous studies show, for instance, that online user ratings can positively impact box office earnings (Chintagunta et al., 2010)

movies is larger for less popular movies or movies from lesser-known studios. However, if controlling for the location of a title on the platform, this effect vanishes.

While theory predicts that under sequential distribution, introducing the higher-valued version maximises profits (Eliashberg et al., 2006; Moorthy, 1984), I observe that TELCO introduces most titles as tvod first, which is the lower-valued version. I am not aware of studies investigating whether the sequential distribution of movies to lease and buy in digital channels is optimal or in which order they should be released. Historically, home video market titles were first available to rent due to video cassettes having a high price. When prices decreased due to technological advances, buying became a viable option for consumers (Ulin, 2013). However, I do not observe such a constraint for digital copies. One reason for TELCO's decision to introduce tvod versions first could be the licensing (Smith and Telang, 2016a; Whitten, 2020), with studios releasing titles to lease first. Alternatively, different from what theory predicts, it could be more profitable for the provider first to offer a leasing version. However, I observe from the data that the proportion of movies introduced as pvod first increases over time, meaning that the provider converges towards what theory would predict: introducing the pvod version, as the higher-valued version, first.

7 Limitations

I am not aware of TELCO's decision-making process on adding a pvod version. It is thus possible that movies becoming available as a second version are different to movies which are only ever available as one version, limiting the explanatory power of the analysis. A randomised experiment would help establishing causality. Since such a setup is not within the scope of my thesis, I use a dummy variable regression to control for time-invariant factors that could influence TELCO's decision to add a second version. Further, I use a propensity score matching method to create a similar control group to the group of movies becoming available as two versions based on observable characteristics. However, PSM is no perfect cure, either: it only allows for creating a control group based on observable characteristics of the movies. Hence, covariates that cannot be observed but which might still vary across control and treatment groups cannot be accounted for within the matching. However this is only a minor concern since I also assume that by increasing similarity among observable characteristics within the groups of movies I also increase similarity among unobservable characteristics. Additionally, Figure 8 shows that the control and treatment group's pre-trends as well as their means differ. The control group is thus not a perfect match to the treatment group. Furthermore, due to data limitations, I run the PSM on a period for which I do not have data on the location of a title within the Video-on-Demand platform menu. Therefore, I cannot include a location control within the PSM regression models, limiting the results since I show in other regression models that movies with a more visible location on the platform benefit from higher monthly sales, on average. Nevertheless, the PSM indicates that adding a version to buy impacts monthly unit sales of the pre-existing version to

lease, which is in line with my main regression results.

The generalisability of the findings is limited since I only consider the case of the addition of a pvod version to a pre-existing tvod version, neglecting those movies that enter the catalogue as a pvod version and are subsequently introduced as a tvod version.

From a managerial point of view, rather than the effect on one version's unit sales from adding the second version, the overall effect on revenue from offering one or two versions of a movie is prevalent. To measure the implications on revenue, I would need to conduct a randomised experiment. I would have one group of movies for which only one version would ever be introduced to the platform and another group of similar movies, for which two versions would be available. Then, I could show whether it is more beneficial to offer one or two versions of a title over time. However, since I do not have access to this type of data, I limit my analysis to the one presented, notwithstanding that I cannot establish the contrast necessary to show the overall effects on revenue. However, I note that I do control for time within my analysis and that I do show that over time, the firm changes its strategy from introducing the tvod version first to introducing the pvod version first.

8 Conclusion

This thesis analyses how the availability of a version of a product in a digital channel impacts sales in another digital channel in the context of a Video-on-Demand platform offering titles to lease and buy, respectively. The results lend support for managerial decision-making since platform providers must decide whether, when and where to include different versions of a movie title to maximise profits. From an academic perspective, the thesis adds to the growing body of literature on digital cross-channel effects by focusing on the interplay of digital buying and digital leasing channels.

To estimate the effect, I use secondary data from a platform provider offering versions of the same title to lease and subsequently to buy. I limit the analysis on titles available first to lease. Depending on the delay between the introduction of the first and second version of a title, adding the purchasable version seems to either have a positive information spillover or a cannibalisation effect. If movies are in the catalogue 13 months and more before becoming available as two versions, adding the second version increases sales of the pre-existing version by 34.44%, on average. I attribute this to an increase in awareness of platform users for the title from adding the second version. For movies becoming available as a second version after a year and less in the catalogue, sales of the pre-existing version decrease upon the introduction of the second version by 55.47%, on average. Movies added more recently to the platform do not seem to benefit from the increase in consumer awareness. Possibly, a cannibalisation effect for the pre-existing version predominates for those titles, with consumers deferring to the purchasable version upon it's availability.

Since the positive impact from adding the second version does not seem to prevail for all

movies and since the analysis suggests a potential cannibalisation effect for recently added titles, the platform provider should consider refraining from introducing a second version to recently added titles. Furthermore, the results suggests that movies more easy to find on the platform sell more on average. Therefore, instead of adding a second version, the provider should consider moving movies to lease introduced only recently at more prominent locations to increase sales. If interested in further understanding the effects accompanying the Video-on-Demand platform provider's versioning strategies and to maximise profits, it would be helpful to run randomised experiments to test their strategies.

Additionally studying the effect of first introducing a title for sale and later for lease would allow a more complete understanding of potential cannibalisation effects. Such an analysis would be especially revealing since theory predicts that in the case of subsequent distribution, adding the higher-priced version first (in this context a version to buy) maximises profits. Furthermore, it remains to be seen whether the simultaneous, instead of the subsequent, introduction would be profit-maximising.

References

Bhargava, H. K. and Choudhary, V. (2008). When is versioning optimal for information goods? *Management Science*, 54(5):1029–1035.

Blaug, M. (1997). *Economic theory in retrospect*. Cambridge University Press.

Calzada, J. and Valletti, T. M. (2012). Intertemporal movie distribution: Versioning when customers can buy both versions. *Marketing Science*, 31(4):649–667.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Chintagunta, P. K., Gopinath, S., and Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, 29(5):944–957.

Danaher, B., Dhanasobhon, S., Smith, M. D., and Telang, R. (2010). Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and internet piracy. *Marketing Science*, 29(6):1138–1151.

Eliashberg, J., Elberse, A., and Leenders, M. A. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Science*, 25(6):638–661.

Frank, B. (1994). Optimal timing of movie releases in ancillary markets: The case of video releases. *Journal of Cultural Economics*, 18(2):125–133.

Gong, J., Smith, M. D., and Telang, R. (2015). Substitution or promotion? The impact of price discounts on cross-channel sales of digital movies. *Journal of Retailing*, 91(2):343–357.

Hennig-Thurau, T., Henning, V., Sattler, H., Eggers, F., and Houston, M. B. (2007). The last picture show? Timing and order of movie distribution channels. *Journal of Marketing*, 71(4):63–83.

Knox, G. and Eliashberg, J. (2009). The consumer's rent vs. buy decision in the rentailer. *International Journal of Research in Marketing*, 26(2):125–135.

Kumar, A., Smith, M. D., and Telang, R. (2014). Information discovery and the long tail of motion picture content. *MIS Quarterly*, 38(4):1057–1078.

Lehmann, D. R. and Weinberg, C. B. (2000). Sales through sequential distribution channels: An application to movies and videos. *Journal of Marketing*, 64(3):18–33.

Moorthy, S. K. (1984). Market segmentation, self-selection, and product line design. *Marketing Science*, 3(4):288–307.

Moorthy, S. K. and Png, I. P. (1992). Market segmentation, cannibalization, and the timing of product introductions. *Management Science*, 38(3):345–359.

Mukherjee, A. and Kadiyali, V. (2011). Modeling multichannel home video demand in the US motion picture industry. *Journal of Marketing Research*, 48(6):985–995.

Shapiro, C. and Varian, H. R. (1998). Versioning: the smart way to sell information. *Harvard Business Review*, 107(6):107.

Smith, M. D. and Telang, R. (2009). Competing with free: The impact of movie broadcasts on DVD sales and Internet piracy. *MIS Quarterly*, pages 321–338.

Smith, M. D. and Telang, R. (2016a). *Streaming, Sharing, Stealing: Big Data and the Future of Entertainment*. MIT Press.

Smith, M. D. and Telang, R. (2016b). Windows of opportunity: the impact of piracy and delayed international availability on DVD sales. Working paper, Heinz College, Carnegie Mellon University.

Swisher, K. (2020). The Window on New Movie Releases Finally Shatters. *The New York Times*. Retrieved from https://www.nytimes.com/2020/12/04/opinion/warner-hbo-movies-theaters.html. Last visited: 2020-12-08.

Ulin, J. (2013). *The business of media distribution: Monetizing film, TV and video content in an online world*. CRC Press.

Varian, H. R. (1995). Pricing information goods. Working paper, Department of Economics, University of Michigan.

Varian, H. R. (1997). Versioning information goods. Working paper, University of California, Berkeley.

Vogel, H. L. (2020). *Entertainment industry economics: A guide for financial analysis*. Cambridge University Press.

Waldfogel, J. (2016). Cinematic explosion: New products, unpredictabilty and realized quality in the digital era. *The Journal of Industrial Economics*, 64(4):755–772.

Weinberg, C. B. (2005). Profits out of the Picture: Research Issues and Revenue Sources Beyond the North American Box Office. In Moul, C. C., editor, *A Concise Handbook of Movie Industry Economics*, page 163–198. Cambridge University Press.

Whitten, S. (2020). Why Hollywood is sticking with movie theaters and only a few films are heading to streaming. *CNBC*. Retrieved from https://www.cnbc.com/2020/04/23/why-hollywood-is-sticking-with-movie-theaters-and-not-streaming.html. Last visited: 2021-01-26.

A Appendix

A.1 Figures

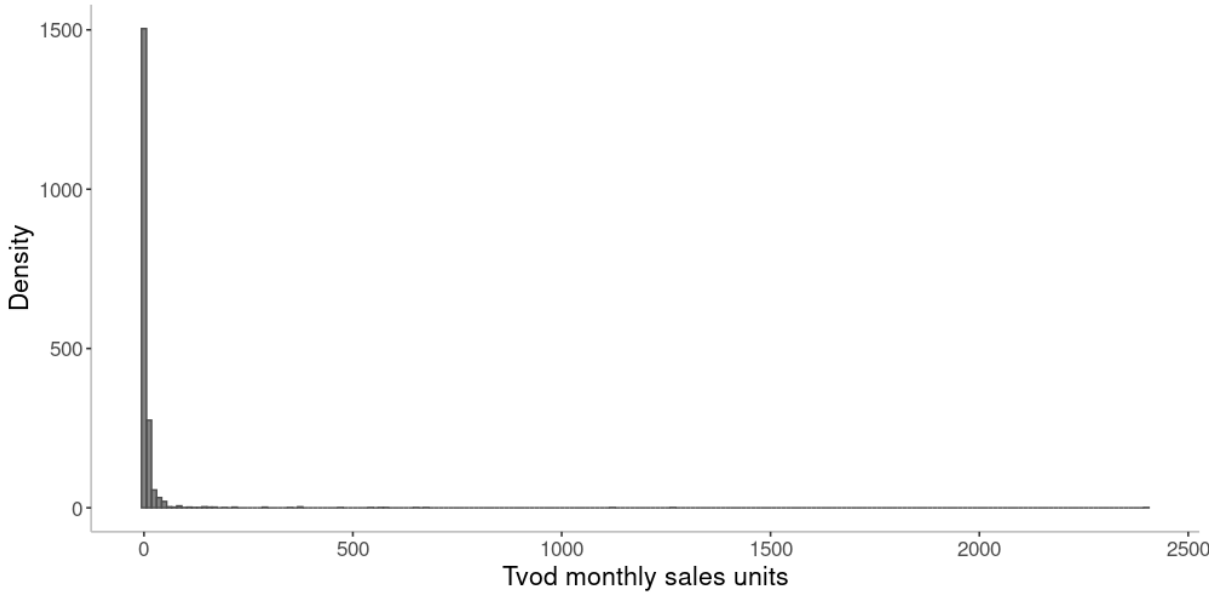

Figure A1: Tvod monthly sales distribution for titles eventually available as both versions

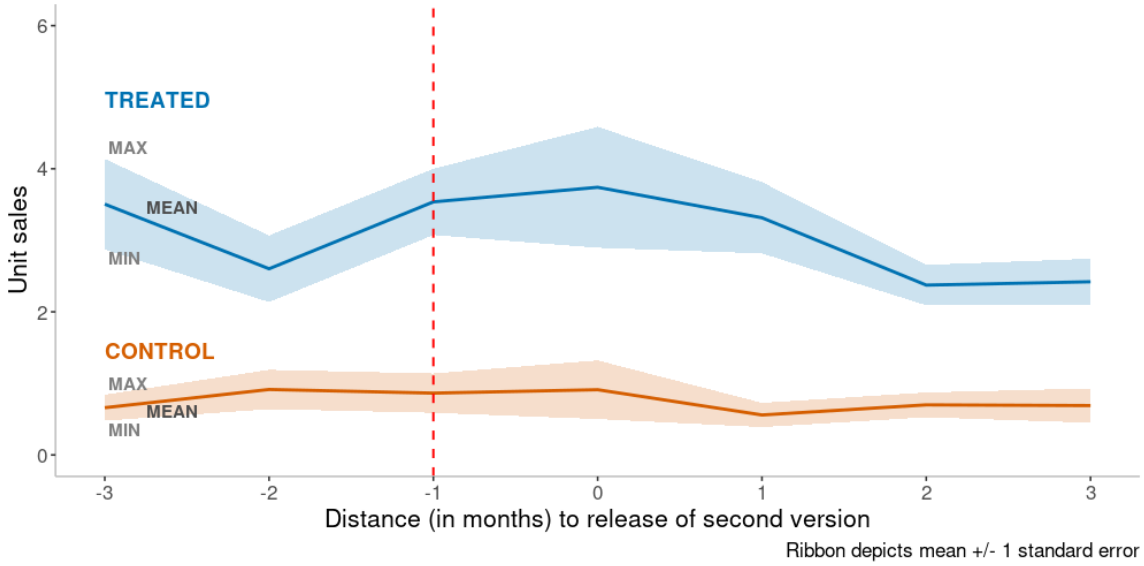

Figure A2: Tvod monthly sales trend - control and treatment group (PSM) using IMDb rating covariate

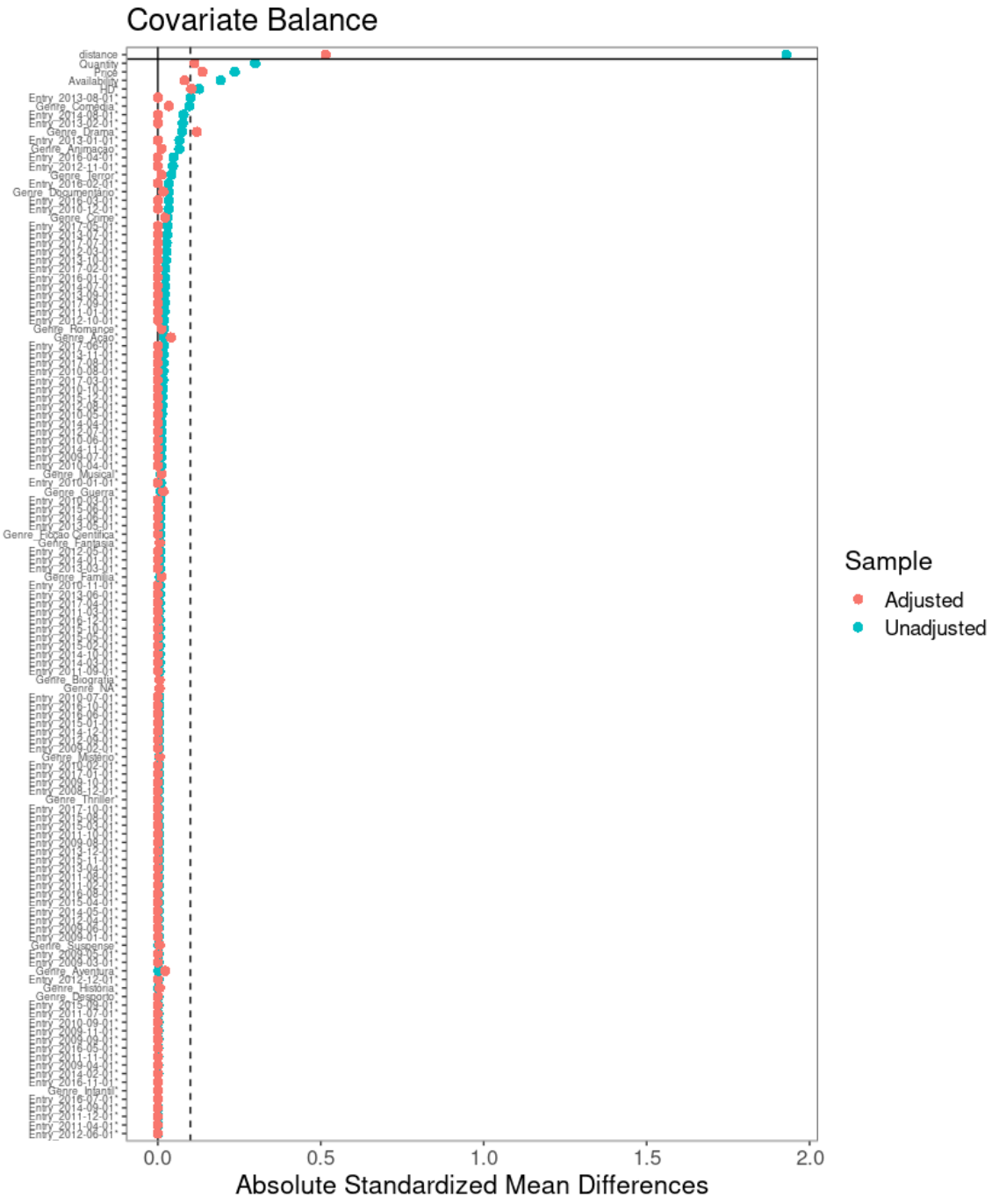

Figure A3: Covariate balance - without IMDb rating

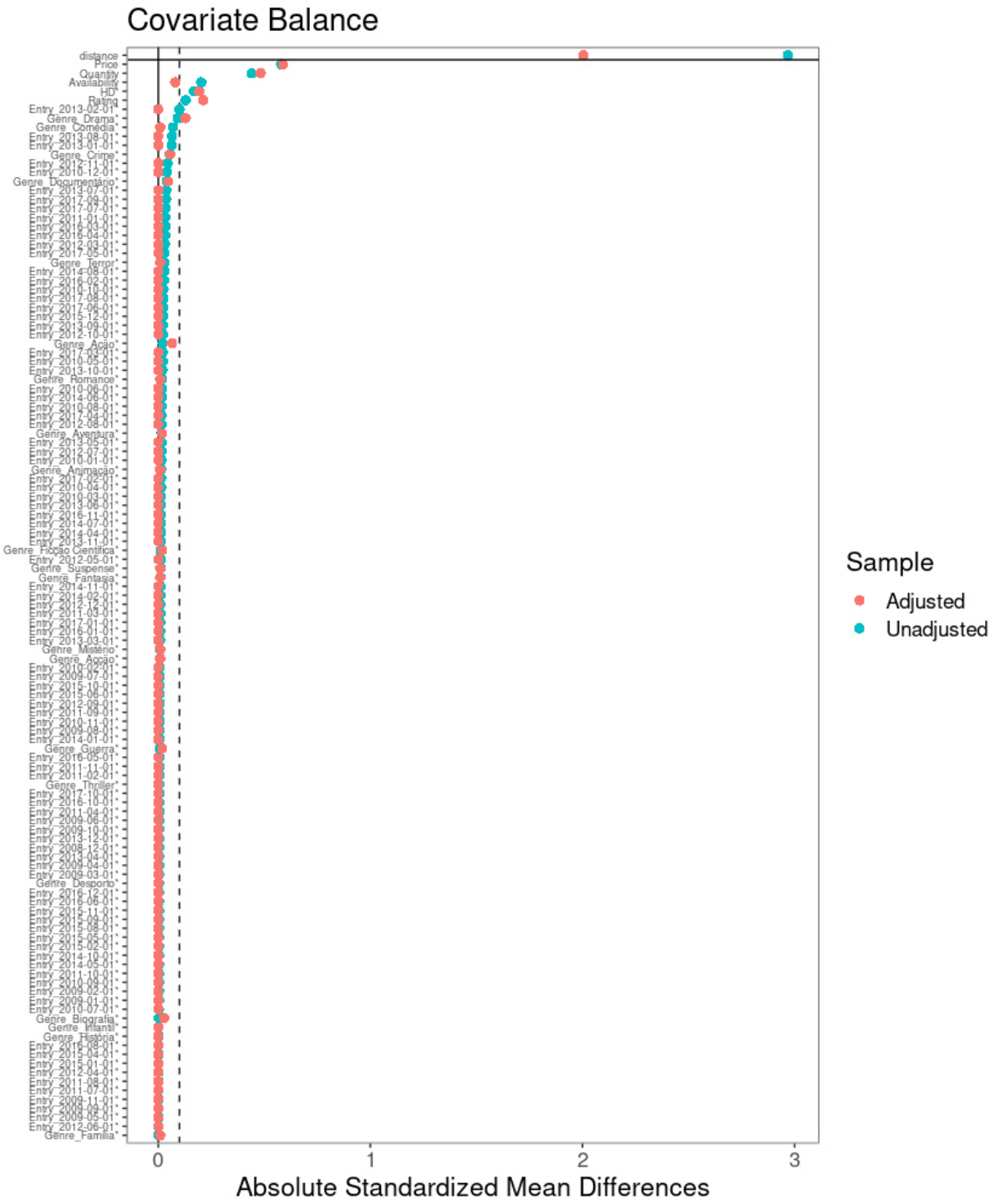

Figure A4: Covariate balance - with IMDb rating

A.2 Tables

Table A1: Average monthly unit sales and prices - titles available as two versions

	all	tvod, one version	tvod, two versions	pvod, one version	pvod, two versions
Titles					
number	503	338	338	165	165
Quantity Sold					
min	0	0	0	0	0
max	3441	3441	1268	873	227
mean	5.78	8.74	3.04	6.81	3.38
sd	41.61	56.72	20.56	40.37	14.93
Price					
min	0.99	0.99	1.49	5.99	5
max	19.99	4.49	4.49	19.99	19.99
mean	4.11	3.2	3.21	14.96	14.78
sd	3.32	0.77	0.65	2.76	3.19

Table A2: Availability as two versions interacted with IMDb rating

	Dependent variable:					
	monthly quantity					
	(1)	(2)	(3)	(4)	(5)	(6)
two versions	1.028*	1.048*	1.122**	0.991	0.278	0.200
	(0.556)	(0.556)	(0.548)	(1.240)	(1.277)	(1.241)
price		−0.865**	−0.807**			−1.705***
		(0.343)	(0.384)			(0.492)
age catalogue			−0.168***			−0.106
			(0.048)			(0.124)
age catalogue^2			0.001**			0.001
			(0.0004)			(0.001)
location					−0.030***	−0.030***
					(0.009)	(0.010)
two versions x IMDb rating	−0.122	−0.130	−0.140*	−0.083	−0.035	−0.028
	(0.084)	(0.084)	(0.082)	(0.171)	(0.178)	(0.173)
Movie ID Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,134	1,134	1,134	284	284	284
Log Likelihood	−2,192.910	−2,190.254	−2,184.545	−585.000	−575.571	−572.613
θ	2.720*** (0.219)	2.741*** (0.221)	2.789*** (0.225)	2.460*** (0.367)	2.790*** (0.429)	2.981*** (0.471)
Akaike Inf. Crit.	4,755.820	4,752.508	4,743.091	1,289.999	1,273.142	1,271.225

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster robust standard errors in ()
Standard errors clustered by movie ID
Including dummy variable per Movie ID, Year and Month

Table A3: Propensity score matching regression results

	Dependent variable:	
	monthly quantity	
	PSM w/o IMDb rating	PSM w/ IMDb rating
	(1)	(2)
two versions	0.398^{***}	0.463^{**}
	(0.132)	(0.207)
price	-0.624^{*}	-0.041
	(0.325)	(0.520)
age catalogue	-0.171^{***}	-0.163^{***}
	(0.037)	(0.052)
age catalogue^2	0.001^{**}	0.001
	(0.0003)	(0.0004)
Movie ID Dummies	Yes	Yes
Year Dummies	Yes	Yes
Month Dummies	Yes	Yes
Observations	2,329	1,460
Log Likelihood	$-3,223.787$	$-1,782.503$
θ	4.773^{***} (0.395)	4.863^{***} (0.590)
Akaike Inf. Crit.	7,179.574	4,035.006

Note: *p<0.1; **p<0.05; ***p<0.01
Cluster Robust Standard Errors in ()
Standard Errors clustered by Movie ID
Including dummy variable per Movie ID, Year and Month