The Impact of Popularity on the Sales of Movies in Video-on-Demand: a Randomized Experiment*

Miguel Godinho de Matos[1,2], Pedro Ferreira[2,3], Rahul Telang[3], and Michael Smith[3]

[1]Instituto Superior Tecnico, Technical University of Lisbon
[2]Department of Engineering and Public Policy, Carnegie Mellon University
[3]Heinz College, Carnegie Mellon University
{*miguelgodinhomatos,pedrof,rtelang,mds*}*@cmu.edu*

Abstract

In this paper, we design and implement a randomized experiment to determine the role that popularity plays on the sales of movies over VoD. We use the VoD system of a large telecommunications provider during half a year in 2012. The popularity of a movie in this system is encoded by the slot in which the movie is displayed on the TV screen. Movies with more likes are shown farthest to the left. During our the experiment, movies were primarily placed in their true slot and shown along with their true number of likes. At random moments, some movies were swapped and thus displayed our of order and with a fake number of likes. The movies that were swapped were selected at random. We find that promoting a movie by one slot increases weekly sales by 4% on average. We find that a movie promoted (demoted) to a fake slot sells 15.9% less (27.7% more) than a true movie placed at that slot, on average across all manipulations we introduced. We show that this asymmetry is related to the amount of information publicly available about the movies manipulated. More well known movies are less sensitive to manipulations. We also find that a movie promoted (demoted) to a fake slot receives 33.1% fewer (30.1% more) likes than a true movie at that slot. Therefore, manipulated movies tend to move back to their true slot over time. Hence, self-fulfilling prophecies are hard to sustain in a market in which goods are costly and sufficiently well known. During this adjustment process, the provider enjoys increased profits while subscribers lose welfare. This process is likely to converge quickly, which might lead the provider to promote different movies over time.

1 Introduction

Figure 1 shows that home video revenues have increased substantially since the 1970s while theater revenues have remained constant over time. One can still argue that the success of a movie dependents highly on box office sales because exhibition in theaters not only allows for covering a significant part of the cost to produce a movie but

*This paper has been prepared for the 2013 Winter Workshop on the Economics of Digitization organized by the National Bureau of Economic Research. This work has been partially supported by the Portuguese Foundation For Science and Technology through the Carnegie Mellon Portugal Program under Grant SFRH/BD/33772/2009.

also triggers demand in subsequent channels. However, it is clear that digitization is changing the structure of the industry. In particular, the share of Video-on-Demand (VoD) and Pay-Per-View (PPV) in the electronic spending on movie rentals in the US increased roughly 4 times between 2000 and 2009. Brick and mortar's share reduced roughly 50% during the same period of time (Waterman, 2011).

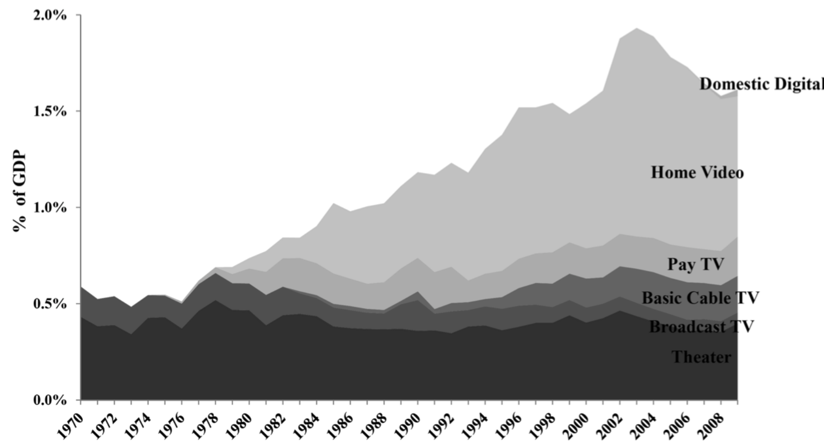


Figure 1: Revenues of movie distributors in the US market as a percentage of GDP (excluding merchandising). Source: (Waterman, 2011)

VoD providers such as Amazon, Netflix or Hulu have catalogs with more than 100,000 titles (Rowinski, 2011), whereas traditional brick and mortar stores offer catalogs with no more than 3,000 titles (Anderson, 2006). Economic theory predicts that product variety increases consumer welfare (Hotelling, 1929; Dixit and Stiglitz, 1977; Salop, 1979). However, search costs also increase with the number of products that consumers need to scan. Therefore, consumers may be unable to internalize the benefits of increased variety (Nelson, 1970; Sawhney and Eliashberg, 1996). In fact, a number of studies have reported a negative relationship between product variety and sales. For example, (Iyengar and Lepper, 2000) showed that increasing the variety of flavors of a specific jam product in a supermarket reduced consumer willingness to buy. (Boatwright and Nunes, 2001) showed that reducing the number of stock keeping units in a grocery store had a positive impact on sales. More recently, (Kuksov and Villas-Boas, 2010) developed a theoretical model that shows that excess variety increases consumer search costs and reduces total sales.

Product variety can increase consumer welfare if more efficient search mechanisms become available. This is particularly true in the movie industry. Several surveys in the US show that consumers welcome recommendations on which movies to watch (De Vriendt et al., 2011), probably because movies are an example of an experience good (Nelson, 1970), (Sawhney and Eliashberg, 1996): their quality can only be ascertain after consumption. 45% of the people surveyed by Ovum in 9 countries around the world welcomed suggestions from friends when searching for new movies to watch (Little, 2010). Tapping into this opportunity, several companies are now implementing recommender systems to provide suggestions to their clients. Again, Hulu, Netflix and Amanzon are widely known examples. These companies incorporate rating mechanisms in their recommender systems whereby consumers are allowed to express whether they liked the content they purchased.

Determining the true impact of rating systems on sales is a challenging empirical question. Observational studies are often subject to the *reflection problem* (Manski, 1993), which hampers the identification of the impact of group behavior on individual decisions. As such, many observational studies offer conflicting perspectives. For example, (Eliashberg and Shugan, 1997) concludes that ratings from movie critics are not good predictors of sales, whereas (Reinstein and Snyder, 2005) concludes otherwise. Several authors used experiments to try to obtain identification. For example, (Salganik et al., 2006) studied the effect of popularity in a market of songs from obscure bands. (Tucker and Zhang, 2011) studied the effect of popularity across wedding service vendors. These studies show that popularity can be, to a certain extent, self-reinforcing. However, they do not explicitly control for the quality of the experience obtained by consumers.

In this paper, we design a randomized experiment to determine the role that social signals play on the sales of VoD products. We use the VoD system of a large telecommunications provider (at which subscribers need to pay to lease movies). Our experiment run live for half a year during 2012. The popularity of a movie in the VoD system of this provider is encoded by the order in which the movie is displayed on the TV screen, hereinafter called the rank, which is a function of the number of likes issued by subscribers. A movie with a higher number of likes is shown farther to the left on the TV screen. During our the experiment, movies were primarily placed in their true rank and shown along with their true number of likes. At random moments, some movies were swapped and thus displayed ”our of order” and with a fake number of likes. The movies swapped were randomly selected. These random exogenous shocks allow for disentangling the perceived quality from the true quality of a movie, thus allowing us to obtain unbiased estimates for the effect of popularity on VoD sales.

We find that on average weekly sales increase by 4% when a movie is promoted one rank. We also find that the weekly sales of a movie promoted (demoted) to a better (worse) rank are 15.9% lower (27.7% higher) than those of a movie placed at that rank by the number of likes issued by subscribers, on average across all manipulations we introduced. We show that this asymmetry is related to the amount of information publicly available about the movies manipulated, as measured by number of IMDb votes. More well known movies are less sensitive to our manipulations. We also find that a movie promoted (demoted) to a better (worse) rank receives 33.1% fewer (30.1% more) likes than a movie placed at that rank by the number of likes issued by subscribers. Therefore, manipulated movies tend to move back to their true rank over time. This means that self-fulfilling prophecies are hard to sustain in a market in which goods are costly and sufficiently well known. Finally, we provide evidence that during this process of adjustment, the provider enjoys increased profits while subscribers lose welfare, as measured by the number of likes issused. This process is likely to converge quickly, which might lead the provider to promote different movies over time.

2 Related Literature

Most papers looking at the impact of quality signals in the movie industry are observational and offer contradictory perspectives. (Litman, 1983) and (Wallace et al., 1993) analyzed 125 and 1687 movies, respectively, released in the US between 1972-78 and 1956-88, respectively. Both report a positive correlation between box office sales and reviews by movie critics. However, (Eliashberg and Shugan, 1997) found that ratings from movie critics are not good predictors of sales during the opening week. They argue that despite being correlated with cumulative movie sales, these ratings do not influence sales in a causal sense.

(Godes and Mayzlin, 2004) studied 44 TV shows released in the US between 1999 and 2000. They found that the dispersion in Word-of-Mouth (WoM) about these shows across distinct groups in Usenet (a news aggregator) was positively correlated to their ratings. However, they were unable to establish a link between WoM, measured by number of conversations about a show, and future rankings, which correlate to sales. (Liu, 2006) studied data from message boards at Yahoo Movies! about 40 movies released between May and September 2002 in the US. They found that the volume of WoM was positively correlated with box office sales but they could not establish a statistically significant relationship between the direction implied in the messages (positive/negative comments) and sales.

A number of previous studies fail to account for the potential correlation between unobserved quality and ratings and therefore are unable to investigate the causal mechanisms that might be at the root of the impact of reviews on sales. Other papers have attempted to overcome this concern. For example, (Reinstein and Snyder, 2005) applied a difference in difference model to a sample of more than 600 movies rated by two influential movie critics to try to identify the marginal impact of reviews on sales. Using the fact that some movie reviews were issued prior to the release of the movie while others were issued after the opening week, they showed that ratings from movie critic were positively correlated with sales and influenced box office sales during the opening week, which again contradicts the findings in (Eliashberg and Shugan, 1997).

(Zhang and Dellarocas, 2006) developed a structural model to study the impact of consumer and movie critic ratings on sales. They showed that good reviews drove movies sales but that the volume and dispersion of the reviews did not. (Dellarocas et al., 2007) developed a predictive model for movie sales that showed that the volume, valence and dispersion of reviews were all positive and statistically significant predictors of box office sales. Finally, (Duan et al., 2008) proposed a model with simultaneous equations to estimate user movie ratings and movie box office sales simultaneously. They concluded that WoM is a strong driver of box office sales, which contradicts the findings in (Zhang and Dellarocas, 2006). Therefore, there is substantial conflict even across the studies that attempt to control for unobserved quality.

A number of authors used experiments to better overcome the traditional hinderances of observational studies. These studies analyze the impact of popularity on sales in the context of other industries. In a seminal paper, (Salganik et al., 2006) created two virtual markets for songs from unknown bands and recruited a group of subjects on a website for teenager interests. Each subject was randomly assigned to one of these markets. Songs were ordered randomly in one of the markets and ordered according to the number of downloads in the other market. Subjects were asked to chose songs to listen, to rate them and then to download them for free if they so wanted. Their study showed that the best (worst) songs received more (less) downloads. The songs in between tended to receive ever more (less) downloads when shown at a higher (lower) rank. In other words, popularity was self-reinforcing for these songs.

In a follow-up study (Salganik and Watts, 2008) run a similar experiment using similar songs and a similar pool of subjects. In a setup phase they ask participants to listen to the songs and to rate them. Then they order songs according to these ratings so that better songs would come last and thus seem worse. In this setting, they observed that over time all songs (good or bad) tended to converge to their true download rank. Taken together, these studies show that self-fulfilling prophecies in these markets are constrained by the individuals' private preferences.

A similar experiment was developed by (Tucker and Zhang, 2011). They used an online hub for online wedding service vendors to explore the impact of popularity on the number of clicks that each vendor obtained. They displayed vendors in three categories. In one category vendors were sorted in decreasing order of the number of clicks received. In another category vendors were sorted in increasing order of the number of clicks received. In both cases, vendors were listed along with the number of clicks received. In the last category vendors were sorted alphabetically and no information on clicks received was displayed. They compared vendors across different categories, before and during their experiment, to determine the impact of popularity, measured by the number of clicks received, on future clicks. They conclude that popularity is self reinforcing and that vendors that operate in narrower markets benefit the most from this dynamics.

Our paper is different from these studies in some important dimensions. First, the papers by (Salganik et al., 2006) and (Tucker and Zhang, 2011) measure impact of *popularity* on sales. They do not measure the impact of user feedback –*like* – on sales. One expects *likes* to reflect better the subscribers' taste and assessment of quality. This is especially true for experience goods like music and movies, for which more downloads typically lead to more *popularity* and vice-versa. In our setting, more *likes* may lead to more purchases. However, the decision to provide *likes* in our case it tightly related to the quality of the movies watched. In short, we believe that *likes* are a better and stronger measure of quality than the *popularity* measures used in previous studies. In (Salganik et al., 2006) downloads might proxy whether subjects like songs but in their settings they are only a noisy measure of preferences across songs.

Another important difference in our setting is that the goods are not free. Subscribers, in our setting, have to make explicit decisions that involve financial risks. The price to rent movies in the VoD system of our Industrial Partner (IP) varied between \$1.30 and \$5.20. In (Salganik et al., 2006) and (Salganik and Watts, 2008) songs could be downloaded for free. Subjects did not incur any financial risk in either listening or downloading a song. (Tucker and Zhang, 2011) observe click through rates on websites but they know nothing about actual purchase decisions. It is not clear how the results of these studies generalize to goods that are not free. For example, in (Salganik and Watts, 2008) demoted songs eventually recover to their true rank. However, this may be an artifact of the fact that songs were provided for free. Since subjects could easily buy several songs, songs in lower ranks may benefit more than demoted movies in our setting.

Another key distinction is that (Salganik et al., 2006) used mostly obscure songs. Thus, downloads provided almost all the information about these songs to the subjects in the study. In most real settings goods are not as unknown to consumers. Consumers can get some information about the quality of products from many external sources. In such settings, the informativeness of *likes* is unclear. We also note that in our setting subjects are real customers of our IP. Our experiment was conducted live in the real field. While this imposes some challenges to carry it out, it also makes for a unique, general and robust setting.

Finally, our paper goes beyond estimating the effect of rank changes on sales. In particular, we are interested in estimating the social cost of changes in rank. Social cost in our context is measured by the loss in sales, or by the the loss in *likes*, when ranks are manipulated. For example, we seek to measure if a movie manipulated into a particular rank sells as much as the correct movie at that rank. Most of the prior work has focused on how rank changes affect sales but not on the social cost associated with these manipulations.

3 The Context of Our Experiment

3.1 The Company and its Dataset

Our experiment was performed using a real world VoD system from a major telecommunications provider, hereinafter called Industrial Partner (IP). Our IP offers TV, Internet, telephone and mobile phone service. IP is the market leader of Pay-TV services in the country where it operates. It services approximately 1.5 million households, 69% of which purchase triple play bundles that include TV, Internet and fixed telephony. According to a market report published by Screen Digest, 65% of the households in this country subscribed to Pay-TV services by the end of 2012. The same report shows that 46% of households with Pay-TV obtained service over cable, 23% over IPTV and the remaining 28% over satellite. Our IP offers Pay-TV through both wired connections and satellite.

We had access to our IP's VoD database between February 2009 and December 2012, which includes information on all of its 3,408,995 subscribers, of which 1,479,895 are on average active at any point in time. 623,516 of the active subscribers subscribe services that include VoD. Overall, 681,036 subscribers watched VoD content at least once and 465,059 subscribers paid for VoD content at least once during this 41-month period. The remaining subscribers with VoD capabilities never used the service. We also had access to all (paid and free of charge) VoD transactions. During this period we observe 89,074,657 transactions, of which 6,293,557 correspond to paid leases.

We have the anonymized identifier of the subscriber requesting each (and every) transaction as well as the anonymized identifier for the MAC address of the specific Set-Top Box (STB) that did so. For each transaction we have a timestamp, the price and the identifier of the movie leased. For each movie in our IP's database we have title, director, studio, play length, synopses, cast, genres, asset type (movie, music, documentary, etc). We also have information on the daily layout of the TV screen that subscribers saw when they logged into the VoD system between 11-2011 and 12-2012. This information includes the tree of menus displayed as well as the order, hereinafter called rank, in which movies were displayed under each menu on the screen from left to right. Menus are associated with different editorial lines as described in the next section. Finally, we also have daily information on all trailer views and on the number of likes issued to each (and every) movie.

3.2 VoD Service and Interface

Our IP provides service over wired and satellite infrastructure. However, satellite subscribers cannot subscribe to VoD. Wired subscribers can obtain one of three types of services: basic/legacy, standard or premium. All of them can watch TV and subscribe to specific channels such as movies, sports, children's and adults' content, etc. As Figure 2 shows, only standard and premium subscribers can use VoD as well as some additional services. For example, both of them can record content if their STB and network connection so allow. Premium subscribers can also restart programs. They can also issue likes for VoD movies and TV programs as well as connect their IP account to Facebook. They are required to also subscribe Internet service. In this paper, we will focus only on standard and premium subscribers. 84% of these subscribers were standard in January 2012. This number reduced to 66% by the end of the year.

Feature	BASIC/LEGACY	STANDARD	PREMIUM
Watch TV	YES	YES	YES
Subscribe to Thematic Channels	YES	YES	YES
Video on Demand	NO	YES	YES
DVR Capabilities	NO	YES	YES
Restart TV Features	NO	NO	YES
Like Ability & Facebook Link	NO	NO	YES
Mandatory Internet Connection Service	NO	NO	YES

Figure 2: Summary of the main features available to subscribers of our IP.

The look and feel of the VoD screen for standard and premium subscribers is different but the organization of content into menus is hierarchically similar. In fact, our IP does not have the ability to suggest different movies to different subscribers, which has a major impact on the way we designed our experiment, as described in the next section. Both standard and premium subscribers can access the VoD system using a hot-key in their STB remote control. When they press it, the entry screen of the VoD system is displayed. This screen, called the *Highlights Section*, contains a set of menus filled with movies, chosen by an editorial team, which are very easy to access. Movies are organized into menus such as Promotions, Suggestions, Newest Releases, etc. Each menu has a header with a name that clearly identifies the type of movies underneath it. Menus are horizontal lines on the screen. Different menus are stacked vertically. Two menus fit on the screen at each time. A cursor highlights a movie cover at a time. Users can scroll across menus. The natural scrolling direction is down, though premium consumers can also scroll up.

Upon scrolling to a new menu, 8 movie covers are visible under that menu and the cursor highlights the movie farthest to the left. Users can also scroll right and left at their leisure within the same menu. Users can scroll right past the last movie cover on the screen to unveil hidden movies under the same menu. There is no limit for the number of movies under a menu though our IP displays no more than 15 movies per menu. The screen of a standard subscriber is somewhat different. Menus show only 4 movies and only 4 other movies can be accessed by scrolling right.

The title and number of likes of the movie highlighted by the cursor are shown on the screen. Standard subscribers do not see the number of likes. Clicking on the cover of a movie leads to a new screen with the year of release, play length, cast, synopsis and number of likes (the latter only in the case of premium subscribers). A number of actions are then available such as lease the movie, use a promotional coupon to lease the movie or watch the movie trailer (if one is available). Premium subscribers can also signal whether they like the movie.

Finally, subscribers can leave the *Highlights Section* of the VoD interface and search for movies in the complete *Catalog* of titles. The catalog is hierarchically organized into content categories such as movies, music, TV-shows, documentaries, etc. Within each of these category screens are organized as described above with menus for genres. Alternatively to browsing through the entire catalog, subscribers can use a keyword search to look for the content of their interest. They can use words that identify titles, movie directors and actors' names.

We note that the likes feature, visible only for premium subscribers, replicates Facebook's well known like button[1]. The number of likes shown along with movie covers is cumulative since the movie's debut at our IP's VoD. Subscribers do not know who liked a particular movie nor who and how many people leased a particular movie.

4 Experimental Design

A new menu, named *The Most Popular During the Past Weeks*, was introduced in the *Highlights Section* of our IP's VoD system. This menu was available for both standard and premium subscribers and included the 15 movies that obtained the highest number of likes in recent times. These movies were shown under the new menu in decreasing order of this number of likes. The experiment run in 1-month cycles for a period of 6 consecutive months. Each cycle was further split into 4 mini-steps of 1 week each[2]. Mini-steps were named true or false. During a true mini-step all movies under this new menu were shown in their true rank. The true number of likes they obtained in the recent past was also shown to premium consumers. During a false mini-step a carefully devised randomization procedure was used to swap some movies under the new menu to separate popularity from unobserved perceived movie quality.

Formally, the experiment ran as follows. Let t_i represent the time at which cycle i began, with $i \in \{1, 2, 3, 4, 5, 6\}$. Let x represent a week's time. At time t_i, we sorted all movies in our IP's VoD system according to the number of likes they received between time $t_i - 2x$ and time t_i. From this list we erased all movies that our IP decided to use in other menus under the *Highlights Section*[3]. We kept the 45 movies

[1] Premium subscribers can also notify IP that they dislike a movie but the number of dislikes is not shown.

[2] Each mini-step started at a time of low VoD usage (Tuesday 2pm).

[3] Our IP decided to list some of these movies under other menus such as Promotions and Suggestions. Cleaning them from our lists avoided listing them under more than one menu in the *Highlights Section*, which would notoriously reduce their search cost. Furthermore, our IP's log system does not allow for identifying the menu under the *Highlights Section* from which a lease originates and thus this cleaning procedure also allows us to ensure that, in fact, leases of movies under the new menu came only from the new menu.

at the head of the resulting list, which we call list L.

After the setup phase described above, which took place at the beginning of each cycle, a true mini-step ensued. Every experimental cycle started with a true mini-step to adjust the subscribers' expectations to the true quality of the movies. We determined the nature of each of the other 3 mini-steps within every cycle using a coin toss[4]. This allowed us to prevent a static pattern of true/false cycles that subscribers could perceive. Table 1 shows the order of true and false cycles used in our experiment.

Table 1: Experimental cycles and nature of the mini-steps

Cycle 1	t_1: True	$t_1 + x$: True	$t_1 + 2x$: False	$t_1 + 3x$: False
Cycle 2	t_2: True	$t_2 + x$: False	$t_2 + 2x$: True	$t_2 + 3x$: False
Cycle 3	t_3: True	$t_3 + x$: False	$t_3 + 2x$: False	$t_3 + 3x$: True
Cycle 4	t_4: True	$t_4 + x$: False	$t_4 + 2x$: False	$t_4 + 3x$: False
Cycle 5	t_5: True	$t_5 + x$: False	$t_5 + 2x$: False	$t_5 + 3x$: False
Cycle 6	t_6: True	$t_6 + x$: False	$t_6 + 2x$: True	$t_6 + 3x$: False

At the beginning of each true mini-step we sorted all movies in L according to the number of likes that they obtain between $t_i - 2x$ and $t_i + nx$ with $n \in \{1, 2, 3\}$ depending on how many mini-steps elapsed since the start of the current cycle. We then displayed under the new menu the first 15 movies in L from left to right on the TV screen. At the beginning of each false mini-step we partitioned list L into 3 sub-lists. List L_1 comprised the 15 movies at the head of list L. List L_2 included the movies between ranks 16 and 30 in list L. Finally, list L_3 contained the movies positioned between ranks 31 and 45 in list L. Then, we performed the following swaps:

- *Within Swap:* we selected y_i and y_j randomly from $\{1, ..., 15\}$ such that $y_i \neq y_j$ and we swapped the number of likes associated with the y_i^{th} and y_j^{th} movies in list L_1;
- *Between Swap:* we selected z_i randomly from $\{1, ..., 15\}$ such that $z_i \neq y_i$ and $z_i \neq y_j$ and we selected z_j randomly from $\{1, ..., 15\}$. Then, we swapped the number of likes of the z_i^{th} movie in list L_1 with the number of likes of the z_j^{th} movie in either list $L2$ or list $L3$, as determined below.

[4]The coin used was biased to ensure a balance between true and false across the whole experience.

The first 15 movies in list L, obtained by concatenating the resulting lists L_1, L_2 and L_3, in this order, were then displayed under the new menu from left to right on the TV screen.

The two types of random swaps introduced with this experiment were aimed at capturing the particular characteristics of the look and feel of the VoD interface. *Within Swaps* allow for determining whether changes in ranks within the list of movies already displayed under the new menu have an impact of sales. *Between Swaps* allow for determining the impact of bringing and removing movies to and from the new menu in the *Highlights Section*. A *Within swap* changes the search cost of the swapped movies only slightly but a *Between Swap* reduces substantially the search costs for a movie that is promoted from the catalog to the new menu and increases substantially the search costs for a movie that is demoted from the new menu into the catalog.

We performed two *Within Swaps* and one *Between Swap* at each false cycle. The latter alternated between lists $L2$ and $L3$, during the first three cycles of our experiment. We performed three *Within Swaps* and two *Between Swaps* at each false cycle, one involving $L2$ and one involving $L3$, during the last three cycles of our experiment. We increased the frequency of swaps in the final three cycles of the experiment to increase the number of treated observations[5].

5 Empirical Model

5.1 Movie Level Specification

The demand for a movie is given by:

$$y_{it} = \lambda + x_{it}\gamma + z_i\theta + w_{it}\tau + c_i + u_{it}, t = 1, ..., T \tag{1}$$

y_{it} is the number of leases of movie i during week t. x_{it} are time varying specific characteristics of movie i such as age, the number of distinct menus where the movie shows up in the VoD catalog and its rank. z_i are time invariant characteristics of

[5]Whether a week is true or false is unrelated to sales during our experiment (results are available upon request).

movie i such as genre, cast and price[6]. w_{it} is the vector of exogenous randomized treatments, which includes promotions and demotions to and from the new menu and our (exogenous random) manipulation of the rank. c_i are unobserved time constant characteristics of movie i such as its story line quality. u_{it} is the idiosyncratic error term.

Equation 1 is the classical fixed effects specification, which we can estimate if we eliminate the unobserved effects in c_i. There is also a potential for serial correlation in the idiosyncratic errors and thus we choose to use robust standard errors and first differences instead of fixed effects to do so (Wooldridge, 2010). Therefore, we estimate the following model:

$$\Delta y_{it} = \beta + \Delta x_{it}\gamma + \Delta w_{it}\tau + \Delta u_{it}, t = 2, ..., T \tag{2}$$

5.2 The Magnitude of Treatment

We must ensure that the magnitude of our exogenous random treatments is appropriately factored into the model in a way that accurately reflects the perceptions of subscribers to these exogenous random shocks. For this purpose, consider two movies, A and B, in ranks a and b, respectively, at time $t_i + nx$, with $n < 3$ under the new menu. When these movies are selected to be swapped, their new ranks in list L are, momentarily, b and a, respectively. At time $t_i + (n + 1)x$, movies in this list are reordered according to number of likes as explained in section 4. As a result, assume that the movie at rank a shifts to rank a' and the movie at rank b shifts to rank b'. Subscribers only see the cumulative effect of swaps and sorting. Thus, from their perspective, movie A moved from rank a to rank b' (therefore a change in rank of $b' - a$) and movie B moved from rank b to rank a' (therefore a change in rank of $a' - b$).

If this swap did not occur, then movie A would have moved from rank a to rank a' and movie B would have moved from rank b to rank b' and this is exactly what sub-

[6]The retail price of each movie takes into consideration the royalty fee paid by our IP and a fixed profit margin per lease. The price of a movie would only change during our experiment if the royalty fees did so, which did not happen. Also, note also that movie prices do not respond to variations in demand at our IP. Our IP does not engage in dynamic pricing. In addition, the network costs for streaming are essentially negligible.

scribers would have seen in this case. Therefore, the effect of the random exogenous swap for movie A is given by $(b' - a) - (a' - a) = b' - a'$. Likewise, the effect of the random exogenous swap for movie B, which is given by $(a' - b) - (b' - b) = a' - b'$. These differences will be introduced into Δw_{it} in equation 1. Note that when a movie is not swapped, this difference is zero. This difference, which hereinafter we will call $rank\_manipulation$, has been coded so that it is positive when a movie is promoted and negative when a movie is demoted. In this example, a' and b' are the true rank for movie A and B, respectively, which hereinafter we will call $rank\_true$. Therefore, we have $rank = rank\_true - rank\_manipulation$.

5.3 Identification and Exogeneity

In our experimental setting, identification is obtained by design. In equation 2, Δw_{it} is not correlated to Δu_{it} because treatments in the former term are exogenous and randomly determined and thus we can safely assume strict exogeneity. A movie is selected to be treated at random. When a movie is selected to be treated it will be swapped with another movie. The latter movie is also selected at random. Therefore, not only a movie is treated at random but also the magnitude of the treatment is random.

Furthermore, Δx_{it} is also not correlated to Δu_{it}. Initially, the number of menus in which a movie shows up in the catalog is determined by which genres the movie is associated with. Changes in the number of menus occur because genres are added or removed or because new categories are editorially created. Additions and removals of genres depend on the variety of genres associated with the set of movies that our IP offers. This set changes over time according to the license windows negotiated with content distributors. For example, for a new genre menu to be created, there has to be a minimum number of movies associated with such genre in this set. These events are not determinants of purchase intentions.

Menus that are created editorially result from the need to aggregate several movies in themes. For example, Christmas specials or Easter specials, etc. There is no reason to believe that editorial choices depend on unobservable determinants of sales other than movie genres. Even if an editorial decision depends on movie quality, the latter is controlled for in c_i. In other words, a potential relationship between number of menus and leases through movie quality is not a problem in our first-difference setting, which controls for the time constant quality of a movie in the fixed effects c_i[7]

Finally, Δx_{it} in equation 2 includes also changes in the rank of movie i. The rank of treated movies under the new menu is determined by the exogenous random swaps as discussed above. The rank of control movies (those that are not swapped) under this menu is determined by the number of likes obtained during the preceding few weeks, as explained in section 4. However, the number of likes and leases must only be related through the quality of the movie. A subscriber is only likely to issue a like for a movie that she leased if the movie was good enough. Recall that subscribers have to pay to watch movies in our setting, which increases the likelihood of truthful likes. But again equation 2 controls for the quality of a movie through the fixed effect c_i.

5.4 Rank Level Specification

Movies are reordered according to the number of likes at the beginning of each mini-step in our experiment. This allows us to take our analysis a step further. This sorting of movies might establish a relationship between rank and the perceived quality of the movie placed at that rank, at least for subscribers of our IP. For example, consider a movie that is promoted k ranks from rank j to rank $j-k$. The models in the previous sub-section measure the impact of such a promotion on the number of leases that this movie obtains. However, they do not evaluate whether this movie sells as much in rank $j-k$ as a movie that had been naturally positioned by subscribers (through the number of likes) at this rank, that is, a true $j-k$ rank movie. The reordering

[7]An external proxy for the quality of a movie is provided by IMDb. IMDb ratings for the movies in IP's VoD system are very stable during our experiment because these movies have already been in theaters before, which must have allowed subscribers to form and stabilize their opinions about the quality of these movies.

of movies at the beginning of each mini-step places the control movies at their true ranks, which allows us to test this hypothesis.

To test whether rank is the sole determinant of movie quality, at they eyes of our IP subscribers, we estimate the following model:

$$y_{rt} = \lambda + x_{rt}\gamma + x_r\theta + w_{rt}\tau + c_r + u_{rt}, t = 1, ..., T \tag{3}$$

y_{rt} represents the leases of the movie at rank r at time t. x_{rt} are characteristics of the movies at rank r at time t, such as price, age and IMDb rating[8]. x_r are time invariant characteristics associated with each rank, which in our setup are extremely unlikely to exist given that movies change ranks frequently. c_r is the intrinsic perceived quality of rank r.

Finally, and as before, w_{rt} is a vector coding our exogenous random treatments, which in this case indicates the type of movie shown at rank r at time t. Treated movies can be from one of two types. A movie can be either demoted to rank r (from a rank higher than r) or promoted to rank r (from a rank lower than r). A movie that is demoted to rank r should have, on average, higher quality than a true rank r movie and, possibly, sell more. The reverse should be true for movies that are promoted. The statistical significance of these dummy variables will reveal how rank affects perceived quality. If they are not statistically significant then rank determines the perceived quality of our IP subscribers. We will estimate equation 3 using a dummy variable for each rank.

[8]Note that the controls for movie quality would drop from the regressions presented before because they are constant over time for the same movie. Here, our units of analysis are ranks and many different movies land in a particular rank over our experiment.

6 Analysis and Results

6.1 Descriptive Statistics

Figure 3 summarizes the evolution of leases in our IP's VoD system over time. This includes the period before and during the experiment. The experiment started on July 3^{rd} 2012. The majority of leases are generated by standard subscribers. Still, premium subscribers are, on average, more intensive users. Both standard and premium users lease more from the highlights section and this bias is stronger for premium subscribers. We observe that overall leases at the *Highlights Section* increased around the time the experiment started, as well as leases from the catalog. The latter, however, declined after a few weeks into the experiment whereas the former did not.

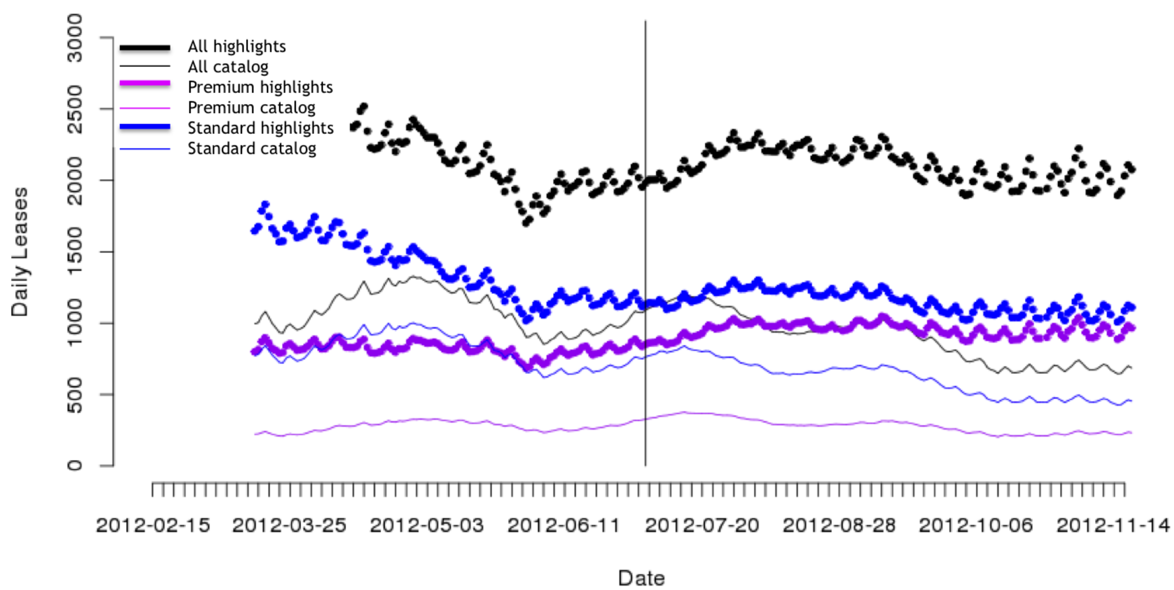
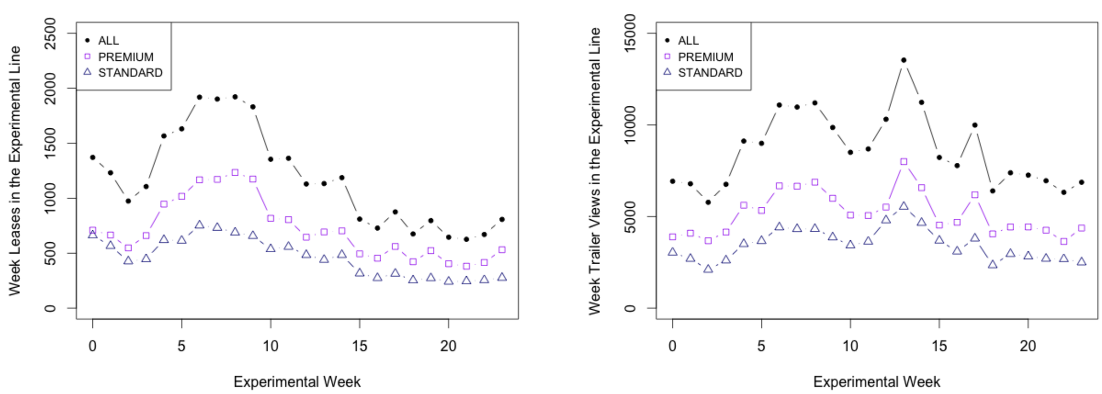


Figure 3: 30-day moving average of daily leases for all movies in IP's VoD system.

Consumption and trailer views in the new menu alone are detailed in panels (a) and (b) of Figure 4, respectively. Unlike the overall VoD usage, the majority of subscribers using the new menu were premium. We note that the new menu is not visible on the entry screen of the VoD system for standard subscribers and they need to hit

10 clicks down to get to it. However, the new menu is readily visible to premium subscribers in the entry screen of the VoD system and accessible with only 1 click up. Furthermore, recall that standard subscribers do not see the number of likes and therefore it might be that the new menu entitled *The Most Popular During the Past Weeks* is not as meaningful to these subscribers.

The new menu started selling well. The number of leases of movies under the new menu has even increased between weeks 2 and 10 of the experiment, which is likely associated with the novelty of the menu. Leases from this menu decreased after week 10 of our experiment. This decline is likely associated to the fact that around week 10 our IP added two additional menus under the *Highlights Section*[9]. These two new menus competed with our experimental menu not only in terms of consumer attention but also in terms of the movies shown. In fact, some of the most liked movies in the past weeks, which by definition would have been pulled into the menu created for this experiment, were in fact used by our IP in these new menus and thus erased from ours to avoid duplication in the *Highlights Section*[10].



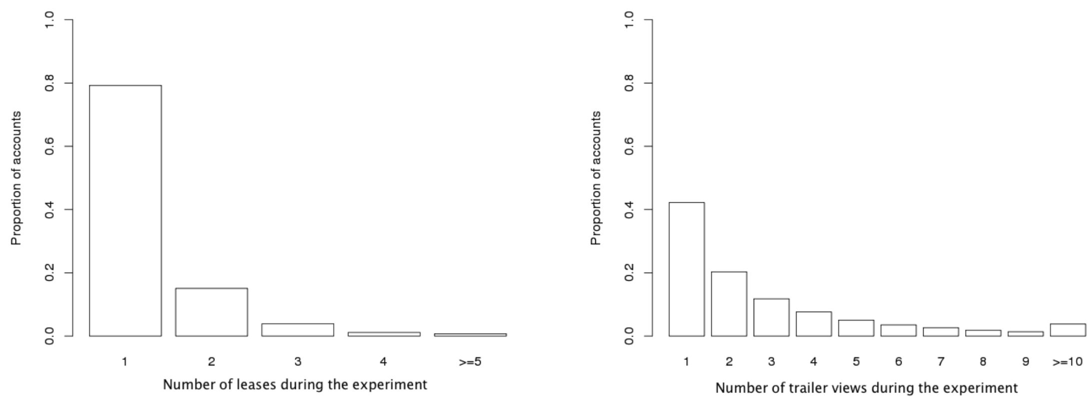
(a) Leases (b) Trailer views

Figure 4: Weekly leases and weekly trailer views for movies in the new menu.

[9]Our IP added two hit lists around this time. One listed the 15 most seen movies of all times. The other listed the most popular movies according to IMDb ratings that were also available in the IP's catalog.

[10]Every month, these two new menus pulled, about 10 to 15 movies that would have been otherwise used in our experiment, half of which would actually have shown in our new menu at the beginning of each month.

Figure 5 shows histograms of behavior per subscriber. 80% of the subscribers leasing movies from the new menu did so only once during the experiment. The distribution of trailer views is relatively more even, which hints to heterogeneity across subscribers in their preferences over movies and in watching trailers to scan for movies to watch. Finally, Table 2 shows the summary statistics for all covariates used in this paper. We show statistics for all movies and separately for treated and control movies in the *Highlight Section* and in the catalog. T-tests to compare the means between control and treatment show that they are statistically similar for all covariates.



(a) Leases

(b) Trailer views

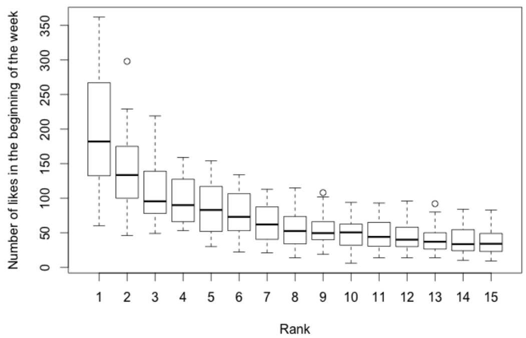
Figure 5: Histograms for leases and trailer views in the new menu during the experiment..

The first panel in Figure 6 shows the number likes as a function of rank at the beginning of each week. By design, this is a decreasing function of rank. We observe however an exponential trend in this relationship. Top movies seem to open a gap in the number of likes to the other movies. The second panel in this figure shows that the number of leases that movies obtain during the week is far from a monotone function of rank, which might suggest that subscribers use more information besides rank to decide which movies to watch. The third panel in this figure shows that movies in better ranks tend to receive more likes during the week. In particular, movies in the visible part of the menu (first 8 ranks for premium subscribers) tend to receive more likes than the movies in the hidden part of the menu. Finally, the fourth panel in this

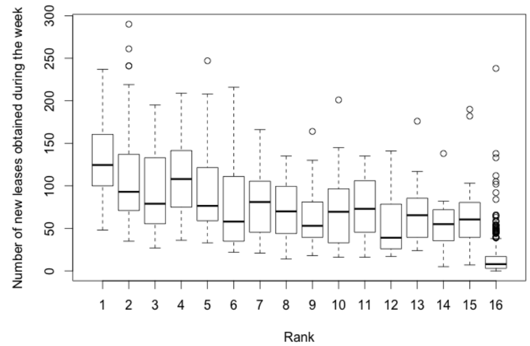
Table 2: Descriptive statistics for the covariates used in this paper.

			Catalog		Highlights	
Vars	Stats	All	Control	Treated	Control	Treated
n_lease						
	mean	36.341	12.85	19.05	80.461	82.29
	sd	45.788	18.201	18.735	50.758	45.904
n_lease_premium						
	mean	19.648	4.174	6.9	48.23	51.527
	sd	29.136	6.124	7.247	33.785	30.23
n_lease_standard						
	mean	16.693	8.676	12.15	32.23	30.763
	sd	19.262	12.644	12.123	20.839	20.25
n_trailerviews						
	mean	241.753	58.56	49.8	573.48	646.333
	sd	321.405	85.345	44.432	290.235	373.671
n_trailerviews_premium						
	mean	134.404	19.417	18	337.25	402.269
	sd	197.865	32.71	13.448	171.831	244.654
n_trailerviews_standard						
	mean	107.349	39.144	31.8	236.23	244.065
	sd	131.237	55.492	32.139	131.944	144.776
rank						
	mean	13.311	16	16	8.531	7.151
	sd	4.487	0	0	4.197	4.369
rank_true						
	mean	13.348	16	8	8.531	9.28
	sd	4.438	0	4.472	4.197	5.247
rank_manipulation						
	mean	0.037	0	-8	0	2.129
	sd	2.5	0	4.472	0	6.811
n_menus						
	mean	1.984	1.708	1.65	2.609	2.258
	sd	1.058	0.932	0.813	1.193	0.674
price						
	mean	287.741	260.883	324	331.617	346.312
	sd	92.662	84.763	96.655	90.21	74.951
imdbrating						
	mean	6.328	6.31	5.98	6.427	6.253
	sd	1.242	1.215	1.485	1.261	1.304
imdbvotes						
	mean	82434.666	87387.516	73270.75	80008.728	58978.022
	sd	114271.836	117701.293	168947.22	111504.825	76944.554
movie_age						
	mean	250.257	291.779	266.294	166.844	187.112
	sd	380.553	415.368	429.838	277.441	314.458
Observations		1017	648	20	256	93

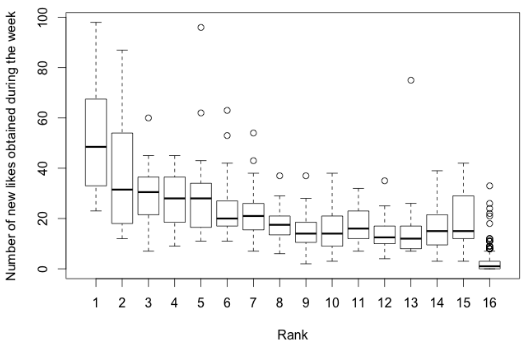
figure shows that on average, at a given rank, promoted (demoted) movies tend to receive fewer (more) likes than true movies. Still, we note that some ranks have very few manipulated movies and thus these averages need to be interpreted with caution.



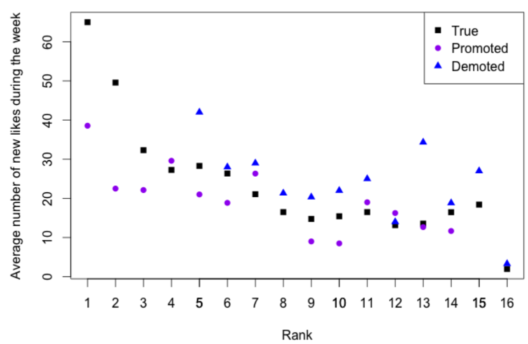
(a) Likes Beginning of the Week



(b) Leases During the Week



(c) Likes During the Week



(d) Likes During the Week

Figure 6: Quartiles and outliers for number of likes and leases.

6.2 The effect of Swaps

We estimate equation 4, which resembles equation 2, to learn the effect of rank on leases. In this regression, $treated\_within * rank\_manipulation$ denotes the size of a rank manipulation within the top 15 ranks, which are the ones shown on the TV screen under the new menu. $promoted\_to\_line$ and $demoted\_from\_line$ denote the size of rank manipulations that lead a movie to go from the catalog into the new menu or to move out from the new menu into the catalog, respectively. These 3 types of manipulations constitute a partition of the space of possible manipulations and therefore their coefficients must be interpreted relative to our control movies. *treated* indicates whether the movie has been treated.

$$
\begin{aligned}
leases_{it} = \lambda + \beta_1 treated_{it} + \beta_2 log(movie\_age_{it}) + \beta_3 n\_menus_{it} + \\
\beta_4 rank\_true_{it} + \beta_5 treated\_within * rank\_manipulation_{it} + \\
\beta_6 promoted\_to\_line_{it} + \beta_7 demoted\_from\_line_{it} + \\
+ week\_dummies + \epsilon_{it}
\end{aligned} \tag{4}
$$

Table 3 shows the results obtained with first-differences for all subscribers and separately for standard and premium subscribers. A movie moved to a better (worse) rank receives more (fewer) leases, as shown by the coefficients on $treated\_within * rank\_manipulation$. This result is statistically significant for both standard and premium subscribers, although less for the former. On average, a manipulation that increases rank by one leads to 2.313 (0.509) more leases by premium (standard) subscribers. This corresponds to an increase of 5.4% (1.6%) in the number of leases. Promoting a movie to the new menu increases 7.2 (2.1) times the leases by premium (standard) subscribers, on average. This significant jump is associated to the difference in search costs between the catalog and the *Highlights Section*. Demoting movies from the new menu yields the opposite effect for premium subscribers. The reduction in the number of leases is 37%. The effect of demotions from the new menu is not statistically significant for standard subscribers. We recall that the new menu was much harder to reach for standard subscribers and thus the standard subscribers that use this menu might already be more willing to search for good movies.

Finally, the coefficient on *treated* is not statistically significant, which indicates that there is no systematic difference in the number of leases obtained by treated and control movies, as expected due to the random assignment of treatments in our experiment.

Table 3: The impact of rank manipulation on leases.

Subscribers	All	Standard	Premium
Model	FD	FD	FD
Variables	$leases_{it}$	$leases_{it}$	$leases_{it}$
(Intercept)	-5.621*	-2.693	-2.928
	(2.892)	(1.479)	(1.89)
	[3.083]	[1.637]	[1.805]
log(movie_age)	-11.852**	-11.775***	-0.076
	(6.389)	(3.268)	(4.176)
	[5.617]	[3.657]	[2.788]
n_menus	12.3***	5.731***	6.569***
	(1.883)	(0.963)	(1.231)
	[3.253]	[1.678]	[1.824]
treated	1.356	0.387	0.969
	(1.857)	(0.95)	(1.214)
	[3.039]	[1.039]	[2.571]
rank_true	-0.62	0.137	-0.756**
	(0.362)	(0.185)	(0.236)
	[0.752]	[0.555]	[0.321]
treated_within * rank_manipulation	2.821***	0.509*	2.313***
	(0.28)	(0.143)	(0.183)
	[0.488]	[0.278]	[0.333]
promoted_to_line	36.31***	9.366***	26.945***
	(3.472)	(1.776)	(2.269)
	[6.091]	[2.084]	[4.579]
demoted_from_line	-23.039***	-4.848	-18.191***
	(3.786)	(1.936)	(2.474)
	[7.091]	[3.131]	[4.646]
$WeekDummies$	Yes	Yes	Yes
N	817	817	817
R-Squared	0.448	0.264	0.478
R-Squared Adj	0.431	0.254	0.461
F-Stat (p-value)	0	0	0

Note 1: Robust standard errors in [ ];
Note 2: *** p<0.01, ** p<0.05, * p<0.1
Note 3: First Differences Estimator

6.3 The Intrinsic Value of Rank

During our experiment, movies were primarily shown in their true rank but sometimes they were exogenously and randomly swapped and thus shown in a fake rank. This variability allows us to study whether a movie placed in a fake rank sells differently from a true movie placed at that rank. A true movie at a rank is a movie that was placed at this rank as a result of the number of likes obtained from subscribers and not as a result of one of our manipulations. We estimate equation 5 that resembles closely equation 3:

$$
\begin{aligned}
leases_{rt} = \lambda + \beta_1 log(movie\_age_{rt}) + \beta_2 n\_menus_{rt} + \beta_3 price_{rt} + \beta_4 imdbrating_{rt} + \\
\beta_5 promoted * treated\_within_{rt} + \beta_6 demoted * treated\_within_{rt} + \\
\beta_7 promoted * treated\_between_{rt} + \beta_7 demoted * treated\_between_{rt} + \\
+ week\_dummies + rank\_dummies + genre\_dummies + year\_release\_dummies + \epsilon_{rt}
\end{aligned} \quad (5)
$$

This regression allows us to compare the number of leases obtained by treated and non-treated movies at each rank. *promoted* (*demoted*) indicates a movie that was promoted (demoted) to a fake rank. *treated_between* indicates whether a rank manipulation entails moving a movie from the catalog to the new menu or vice-versa. Therefore, the 4 types of manipulations included in this regression constitute a partition of the space of possible manipulations and thus their coefficients must be interpreted relative to our control movies.

Table 4 shows the results obtained. The first three columns in this table show the effect of rank manipulations on leases whereas the last column shows the effect of rank manipulations on the number of likes. A movie that is demoted to a fake rank within the new menu sells more than a true movie at that rank. It seems that consumers are able to spot a high quality movie even if it has been shifted to the right on the TV screen under the new menu. This is true for both standard and premium subscribers though less statistically significant for the former. Conversely, a movie that is promoted to a fake rank within the new menu sells less than a true movie at that rank. However, this result is weaker than the effect of demotions within the new menu. Both its magnitude and its statistical significance are lower and, in fact this, effect is not statistically significant for standard subscribers alone.

The asymmetry in the magnitude of these effects indicates that the gain in sales of the demoted movie relative to the sales that the promoted movie would have had at the that rank is enough to counter the loss in sales of the promoted movie relative to the sales that the demoted movie would have had at that rank. Therefore, the provider can increase revenues by strategically manipulating movies across ranks.

Table 4: The effect of promotion and demotion on leases and number of likes at each rank.

	Leases			Likes
Subscribers	All	Standard	Premium	Premium
Variables	$leases_{rt}$	$leases_{rt}$	$leases_{rt}$	$likes_{rt}$
(Intercept)	63.134***	0.986	62.148***	45.615***
	(21.573)	(11.187)	(13.157)	(6.587)
	[18.904]	[10.935]	[10.906]	[5.808]
promoted * treated_within	-12.184*	-2.789	-9.396*	-7.614***
	(4.773)	(2.475)	(2.911)	(1.458)
	[7.308]	[3.221]	[5.108]	[2.584]
demoted * treated_within	22.348***	7.756*	14.592***	6.955***
	(4.749)	(2.463)	(2.896)	(1.45)
	[7.034]	[4.501]	[4.21]	[2.511]
promoted * treated_between	-4.331	1.09	-5.421	-7.853**
	(5.67)	(2.94)	(3.458)	(1.731)
	[8.566]	[3.391]	[5.832]	[3.156]
demoted * treated_between	5.686	4.35*	1.336	1.241
	(5.718)	(2.965)	(3.487)	(1.746)
	[4.448]	[2.451]	[2.524]	[1.051]
log(movie_age)	-3.103***	-1.814***	-1.289**	-0.305
	(0.727)	(0.377)	(0.443)	(0.222)
	[1.178]	[0.689]	[0.633]	[0.275]
n_menus	4.000*	3.891***	0.109	1.276**
	(1.115)	(0.578)	(0.68)	(0.34)
	[2.106]	[1.207]	[1.063]	[0.531]
price	-0.005	-0.006	0.002	0.01*
	(0.016)	(0.008)	(0.01)	(0.005)
	[0.026]	[0.016]	[0.013]	[0.006]
imdbrating	3.541**	1.005	2.536**	1.112*
	(0.759)	(0.394)	(0.463)	(0.232)
	[1.429]	[0.717]	[1.096]	[0.576]
Week Dummies	Yes	Yes	Yes	Yes
Rank Dummies	Yes	Yes	Yes	Yes
Genre Dummies	Yes	Yes	Yes	Yes
Year Release Dummies	Yes	Yes	Yes	Yes
N	1001	1001	1001	1001
R-Squared	0.759	0.631	0.775	0.777
R-Squared Adj	0.697	0.58	0.713	0.714
F-Stat (p-value)	0	0	0	0

Note 1: Robust standard errors in [ ]; **Note 2:** *** p<0.01, ** p<0.05, * p<0.1

The last column of table 4 shows that a movie promoted (demoted) to a fake rank receives fewer (more) likes from subscribers than a true movie at that rank. This result entails that over time, manipulated movies are likely to come back to their true ranks and therefore the provider might only be able to take advantage of manipulations for short periods of time. One possible way to sustain a positive profit out of manipulations is therefore to manipulate movies often and to manipulate different movies over time. However, it is not clear that in the long run subscribers will still rely on our IP's rating system if too many movies are manipulated all the time. Still, our results also show that search costs have a significant impact on sales, therefore, one way for the operator to increase profits is to selectively show and hide movies according to their profit margins. An operator that can do this at the subscriber level can show the most appropriate movies to each subscriber separately (instead of showing to every subscriber the movies that might match average preferences), which is also likely to increase profits without the need to manipulate rankings in complex ways.

Finally, we note that the coefficients for the effects of manipulations interacted with *treated_between* are essentially not statistically significant, which means that search costs dominate the effect of manipulations. If anything, standard subscribers lease demoted movies to worse ranks more than a true movies at these ranks, which again confirms the idea that standard subscribers are more willing to scan for movies to watch.

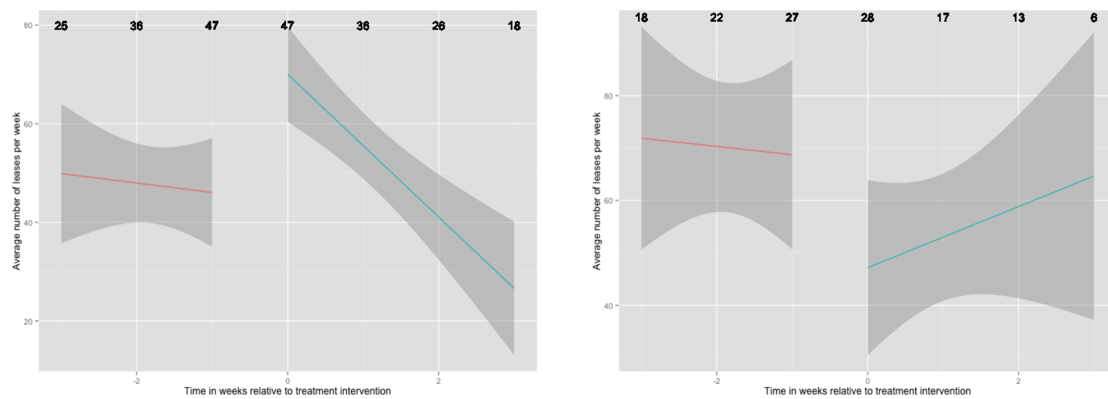
6.4 Convergence and Effect on Subscribers

Our experiment ran in monthly cycles of weekly mini-steps. Movies were refreshed often and some movies were treated more than once (the appendix provides results eliminating sequences of treatments). All these facts combined make it hard to provide a precise estimate for the speed at which the VoD system corrects the exogenous manipulations that we introduced. Yet, Figure 7 depicts a simple descriptive analysis for promoted movies in panel (a) and for demoted movies in panel (b) using data from our experiment. The horizontal axes represent time relative to the moment of treatment. The vertical axes represent, for a particular time t in the horizontal axis, the average number of weekly leases across all movies in our sample that were t days away from their treatment date[11]. On the top of each panel we indicate over how

[11] In fact, in between weeks (days -21, -14, -7, 0, +7, +14, +21) we linearly interpolate weekly sales.

many movies each average is computed[12].

All movies were progressively selling less before treatment due to aging. Promoted (demoted) movies sell significantly more (less) right after treatment. Within 2 weeks promoted movies sell as much as they used to sell before treatment. Demoted movies need about 3 weeks to sell as much as they used to sell before treatment. It seems that demoted movies take longer to climb back to their true rank than promoted movies take to fall back to their true rank. This asymmetry might be related to the fact that the movies in our setting are well known and not for free, but it must also be shaped by the fact that they age quickly with time.



(a) Promoted movies

(b) Demoted movies

Figure 7: The adjustment process on sales after treatment (bands represent 95% CI on leases).

We proceed as follows to assess the impact of the adjustment process described above on subscribers. For each rank, we average the number of weekly likes across movies placed in a true rank and across movies placed in a fake rank. We add up the former and the latter over the entire period of the experiment and we compare them. The sum of the average number of weekly likes is 376.6 across movies placed in true ranks and 355.1 across movies placed in fake ranks. From these statistics, we observe that manipulating movies made subscribers slightly worse off (roughly 5%),

[12]Note that some movies were treated more than once. To avoid confounding from subsequent treatments on the same movie, we only include in this analysis the first treatment of a movie and its data up to a second treatment or up to the end of the panel if the movie was only treated once.

if one uses the number of likes as a measure subscriber well being.

6.5 The Role of Outside Information

We test whether outside information about the movies shown at our IP's VoD system mediates the effects of promoting and demoting movies. We use the number of IMDb votes as a proxy for how well a particular movie is known to consumers in general. The number of IMDb votes of a movie indicates how may people evaluated that movie irrespective of the rating provided. Figure 8 shows that there is significant variation in the number of IMDb votes across movies in our sample. This is not surprising given the well established super star effect in the movie industry, whereby popular movies concentrate a disproportionate amount of attention and therefore are more widely known (Elberse and Oberholzer-Gee, 2006). We hypothesize that the movies in our IP's VoD system that have more outside information are less sensitive to our exogenous random manipulations. This would be consistent with the findings in (Tucker and Zhang, 2011) that show that products with broader appeal are less likely to benefit from increased popularity established in the context of the platforms where they are sold[13].

We classify each movie in our sample according to the number of IMDb votes received until December of 2012. We define a dummy variable called $top25\_imdbvotes$ to indicate wheter a movie is in the top quartile of the distribution of IMDb votes in our sample. We estimate equation 4 adding an interaction term between $rank\_manipulation$ and this new dummy variable. In this regression, this interaction term captures the difference in the effect of our rank manipulations for movies in the top quartile of the distribution of IMDb votes relative to the effect on all the other movies in our sample that were also manipulated. Table 5 presents the results obtained. The effect of the interaction between $rank\_manipulation$ and $top25\_imbdvotes$ is negative and statistically significant thus confirming our hypothesis.

[13](Salganik and Watts, 2008) report a similar result but their measure of appeal is endogenous to the population of subjects used in their experiment.

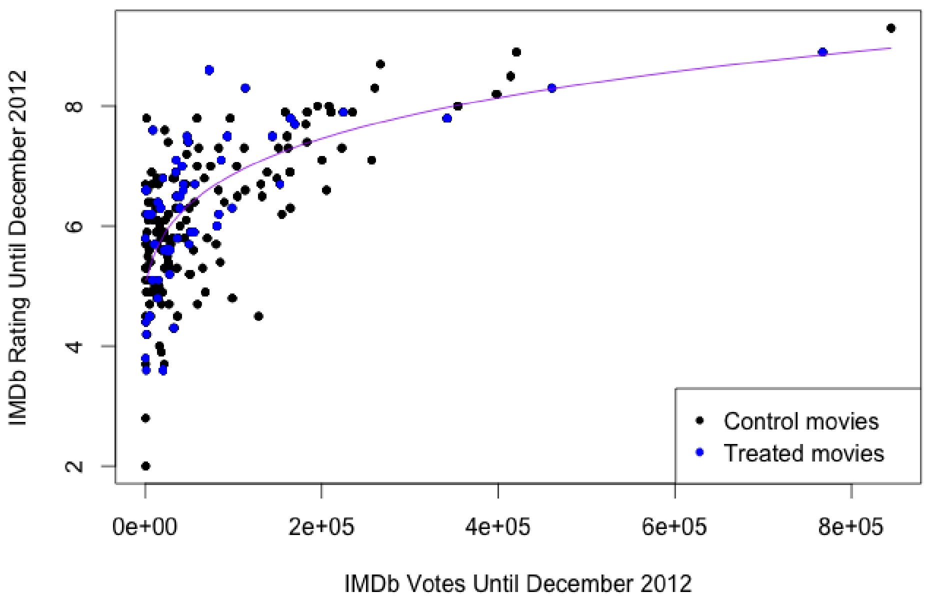


Figure 8: IMDb votes across movies in our sample.

7 Conclusions

In this paper, we design and implement a randomized experiment to determine the role that *likes* play on the sales of movies over VoD. We use the VoD system of a large telecommunications provider during half a year in 2012. A new menu in the *Highlights Section* of this VoD system was introduced showing the most liked movies in the past few weeks. Movies with more *likes* were shown farthest to the left on the TV screen. During our the experiment, movies were primarily placed in their true rank and shown along with their true number of *likes*. At random moments, some movies were swapped and thus displayed our of order and with a fake number of *likes*. The movies that were swapped were selected at random. Randomization allows us to disentangle *likes* from unobserved perceived quality and thus estimate the effect of the former on sales.

We found that search costs play a major role on sales. A movie brought from the catalog into the new menu sells about 7 times more, on average. We found that

Table 5: The role of IMDb votes on the effect of our rank manipulations on leases.

Subscribers	All	Standard	Premium
Model	FD	FD	FD
Variables	$leases_{it}$	$leases_{it}$	$leases_{it}$
(Intercept)	-5.698*	-2.723*	-2.975*
	(2.883)	(1.477)	(1.885)
	[3.092]	[1.647]	[1.802]
log(movie_age)	-11.932**	-11.807***	-0.125
	(6.369)	(3.263)	(4.165)
	[5.624]	[3.665]	[2.78]
n_menus	12.346***	5.749***	6.597***
	(1.877)	(0.962)	(1.228)
	[3.268]	[1.682]	[1.833]
treated	1.193	0.323	0.87
	(1.852)	(0.949)	(1.211)
	[2.901]	[1.037]	[2.479]
rank_true	-0.674	0.115	-0.789**
	(0.361)	(0.185)	(0.236)
	[0.756]	[0.559]	[0.322]
treated_within * rank_manipulation	3.031***	0.591*	2.44***
	(0.292)	(0.149)	(0.191)
	[0.51]	[0.304]	[0.345]
treated_within * rank_manipulation * top25imdbvotes	-2.547**	-1.001*	-1.546**
	(1.037)	(0.531)	(0.678)
	[1.133]	[0.552]	[0.678]
promoted_to_line	36.439***	9.416***	27.023***
	(3.462)	(1.773)	(2.264)
	[6.031]	[2.087]	[4.532]
demoted_from_line	-23.461***	-5.014*	-18.447***
	(3.778)	(1.935)	(2.47)
	[6.736]	[3.025]	[4.439]
$WeekDummies$	Yes	Yes	Yes
N	817	817	817
R-Squared	0.452	0.267	0.482
R-Squared Adj	0.435	0.257	0.463
F-Stat (p-value)	0	0	0

Note 1: Robust standard errors in [ ];
Note 2: *** p<0.01, ** p<0.05, * p<0.1;
Note 3: First Differences Estimator

promoting a movie by one rank increases weekly sales by 4% on average. We found that a movie promoted (demoted) to a fake slot sells 15.9% less (27.7% more) than a true movie placed at that slot, on average across all manipulations we introduced. We showed that this asymmetry is related to the amount of information publicly available about the movies manipulated. More well known movies are less sensitive to manipulations.

We also found that a movie promoted (demoted) to a fake slot receives 33.1%

fewer (30.1% more) likes than a true movie at that slot. Therefore, manipulated movies tend to move back to their true slot over time. During this adjustment process, the provider enjoys increased profits while subscribers lose welfare. This process is likely to converge quickly, in a matter of 2 to 3 weeks time, which might lead the provider to promote different movies over time to sustain its profit margin. However, it is not clear whether in the long run subscribers will believe in the number of *likes* exhibited at this VoD system if movies are manipulated often. Another way for the provider to attract attention to, and possibly increase the sales of, specific movies without manipulating their rank is to strategically show and hide movies between the *Highlights Section* and the catalog.

We have measured the impact of *likes* in a real VoD system of a large telecommunications provider. We believe that number of *likes* is a more truthful measure of the quality experienced by subscribers than several popularity measures previously used in the literature. This is specially true in our setting, in which subscribers need to explicitly make decisions that entail financial risks because movies are not free. Because movies are not free in this setting, demoted movies could be unable to climb back to their true rank. We showed they do at a slower pace than promoted movies fall back to their true rank. The fact that movies are well known in our setting reduces the risk associated with choosing a good but demoted movie. Certainly, trailers also allow subscribers to better perceive the quality of a movie before they pay to watch it, which could benefit demoted movies as long as subscribers are willing to search beyond the first ranked movies.

References

Chris Anderson. *The Long Tail: Why the Future of Business is Selling Less of More.* Hyperion, July 2006. ISBN 1401302378. 1

Peter Boatwright and Joseph C. Nunes. Reducing assortment: An attribute-based approach. *Journal of Marketing*, 65(3):50–63, July 2001. ISSN 0022-2429. URL `http://www.jstor.org/stable/3203466`. ArticleType: research-article / Full publication date: Jul., 2001 / Copyright © 2001 American Marketing Association. 1

Johan De Vriendt, Natalie Degrande, and Marc Verhoeyen. Video content recommendation: An overview and discussion on technologies and business models. *Bell Labs Technical Journal*, 16(2):235–250, September 2011. ISSN 1538-7305. doi: 10.1002/bltj.20513. URL `http://onlinelibrary.wiley.com/doi/10.1002/bltj.20513/abstract`. 1

Chrysanthos Dellarocas, Xiaoquan (Michael) Zhang, and Neveen F. Awad. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21(4):23–45, January 2007. ISSN 10949968. doi: 10.1002/dir.20087. URL `http://linkinghub.elsevier.com/retrieve/pii/S1094996807700361`. 2

Dixit and Stiglitz. Monopolistic competition and optimum product diversity. *The American Economic Review*, Vol. 67(No. 3):297–308, 1977. URL `file:///Users/Miguel.GodinhoMatos/Data/MyFiles@WorkPC/MyLibrary/MyEndNoteLibrary.Data/PDF/[DixitandStiglitz]-(1977paper)-MonopolisticCompetitionandOptimumProductDiv-1755811866/[DixitandStiglitz]-(1977paper)-MonopolisticCompetitionandOptimumProductDiversity.pdf`. 1

Wenjing Duan, Bin Gu, and Andrew B. Whinston. Do online reviews matter? — an empirical investigation of panel data. *Decision Support Systems*, 45(4):1007–1016, November 2008. ISSN 0167-9236. doi: 10.1016/j.dss.2008.04.001. URL `http://www.sciencedirect.com/science/article/pii/S0167923608000754`. 2

Anita Elberse and Felix Oberholzer-Gee. Superstars and underdogs: An examination of the long tail phenomenon in video sales. *Harvard Business School Working Paper Series*, September 2006. 6.5

Jehoshua Eliashberg and Steven M. Shugan. Film critics: Influencers or predictors? *Journal of Marketing*, 61(2):68, April 1997. ISSN 00222429. doi: 10.2307/1251831. URL `http://www.jstor.org/discover/10.2307/1251831?uid=3738880&uid=2129&uid=2&uid=70&uid=4&sid=21101210305387`. 1, 2

David Godes and Dina Mayzlin. Using online conversations to study word-of-mouth communication. *Marketing Science*, 23(4):545–560, 2004. ISSN 0732-2399. URL `http://www.jstor.org/stable/30036688`. ArticleType: research-article / Full publication date: Fall, 2004 / Copyright © 2004 INFORMS. 2

Harold Hotelling. Stability in competition. *The Economic Journal*, 39, No 53:41–57, 1929. 1

S S Iyengar and M R Lepper. When choice is demotivating: can one desire too much of a good thing? *Journal of personality and social psychology*, 79(6):995–1006, December 2000. ISSN 0022-3514. URL `http://www.ncbi.nlm.nih.gov/pubmed/11138768`. PMID: 11138768. 1

Dmitri Kuksov and J. Miguel Villas-Boas. When more alternatives lead to less choice. *Marketing Science*, 29(3):507–524, May 2010. ISSN 0732-2399, 1526-548X. doi: 10.1287/mksc.1090.0535. URL `http://mktsci.journal.informs.org/content/29/3/507`. 1

Barry R. Litman. Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, 16(4):159–175, 1983. ISSN 1540-5931. doi: 10.1111/j.0022-3840.1983.1604_159.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.0022-3840.1983.1604_159.x/abstract`. 2

M. Little. Video discovery and recommendation: Consumer behaviour survey. Technical report, Ovum, January 2010. 1

Yong Liu. Word-of-mouth for movies: Its dynamics and impact on box office revenue. *SSRN eLibrary*, January 2006. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1949819. 2

Charles F. Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993. URL file:///Users/Miguel.GodinhoMatos/Data/MyFiles@WorkPC/MyLibrary/MyEndNoteLibrary.Data/PDF/Manski1993-1506494225/Manski1993.pdf. 1

Phillip Nelson. Information and consumer behavior. *Journal of Political Economy*, 78 (2):311–329, March 1970. ISSN 0022-3808. URL http://www.jstor.org/stable/1830691. ArticleType: research-article / Full publication date: Mar. - Apr., 1970 / Copyright © 1970 The University of Chicago Press. 1

David A. Reinstein and Christopher M. Snyder. THE INFLUENCE OF EXPERT REVIEWS ON CONSUMER DEMAND FOR EXPERIENCE GOODS: a CASE STUDY OF MOVIE CRITICS*. *The Journal of Industrial Economics*, 53(1):27–51, 2005. ISSN 1467-6451. doi: 10.1111/j.0022-1821.2005.00244.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.0022-1821.2005.00244.x/abstract. 1, 2

Dan Rowinski. Amazon's video on demand now has 100,000 titles available. http://www.readwriteweb.com/archives/amazons_video_on_demand_now_has_100000_titles_avai.php, August 2011. URL http://www.readwriteweb.com/archives/amazons_video_on_demand_now_has_100000_titles_avai.php. 1

Matthew J. Salganik and Duncan J. Watts. Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71(4):338–355, December 2008. ISSN 0190-2725, 1939-8999. doi: 10.1177/019027250807100404. URL http://spq.sagepub.com/content/71/4/338. 2, 13

Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*

(New York, N.Y.), 311(5762):854–856, February 2006. ISSN 1095-9203. doi: 10.1126/science.1121066. PMID: 16469928. 1, 2

Steven C. Salop. Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141, 1979. ISSN 0361915X. doi: 10.2307/3003323. URL `http://www.jstor.org/discover/10.2307/3003323?uid=3738880&uid=2&uid=4&sid=21101211562807`. 1

Mohanbir S. Sawhney and Jehoshua Eliashberg. A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, 15(2):113–131, January 1996. ISSN 0732-2399. URL `http://www.jstor.org/stable/184189`. ArticleType: research-article / Full publication date: 1996 / Copyright © 1996 INFORMS. 1

Catherine Tucker and Juanjuan Zhang. How does popularity information affect choices? a field experiment. *Management Science*, 57(5):828–842, May 2011. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.1110.1312. URL `http://mansci.journal.informs.org/content/57/5/828`. 1, 2, 6.5

W. Timothy Wallace, Alan Seigerman, and Morris B. Holbrook. The role of actors and actresses in the success of films: how much is a movie star worth? *Journal of Cultural Economics*, 17(1):1–27, 1993. ISSN 0885-2545. doi: 10.1007/BF00820765. URL `http://www.springerlink.com/content/u08531w4qh160161/abstract/`. 2

David Waterman. Technology and the U.S. movie industry, May 2011. URL `http://is.jrc.ec.europa.eu/pages/ISG/documents/DavidWaterman_sevilla-movies-2011-final.pdf`. 1, 1

Jeffrey M Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, second edition edition, October 2010. ISBN 0262232588. 5.1

Xiaoquan Zhang and Chrysanthos Dellarocas. The lord of the ratings: Is a movie's fate is influenced by reviews? *ICIS 2006 Proceedings*, December 2006. URL `http://aisel.aisnet.org/icis2006/117`. 2

A Additional Analyses

A.1 Impact of Rank on Trailer Views

We replicate the regressions shown in the main paper using trailer views as our dependent variable. In this case, we want to learn whether manipulating the rank of a movie has an effect on the number of trailers watched. Table 6 shows the results obtained, which are qualitatively similar to the ones obtained before for the case of leases[14]. However, both the statistical significance and the magnitude of the impact of manipulations are higher than before. Watching trailers is free of charge (the only resource that consumers commit when watching a trailer is time). It seems that the number of likes attracts consumers to watch trailers and thus likes can be a productive tool to attract consumers to particular movies. This, however, does not necessarily translate into more leases as subscribers do use trailers to form a more certain opinion about the quality of the movies.

B Eliminating Sequences of Treatments

A potential problem with our experiment is the fact that the same movies can be subject to different treatments in consecutive weeks. In this case the effect of the first treatment might contaminate the effect of the second treatment. To assess the impact that such potential contamination might have on our experiment we perform the regressions presented in the main paper but discard all observations of a movie within the same cycle beyond (and including) the second treatment. Because treatment assignment is random, eliminating these observations is equivalent to random attrition in the sample. For each movie that we trim we include a dummy variable indicating whether that movie was trimmed. This dummy variable should not be statistically significant if our assumption of random attrition holds. The trimming operation discards 42 observations (34 treated and 8 after the first treatment).

Table 7 shows the results obtained, which reinforce our previous findings. Manip-

[14] Which could be expected since the correlation between leases and likes is around 0.74 ($p - value < 0.01$)

Table 6: Effect of rank on trailer views.

Subscribers	All	Standard	Premium
Model	FD	FD	FD
Variables	$trailerviews_{it}$	$trailerviews_{it}$	$trailerviews_{it}$
(Intercept)	-12.071	-8.692	-3.379
	(23.021)	(8.895)	(15.748)
	[13.891]	[7.152]	[9.061]
log(movie_age)	-37.535	-55.774**	18.239
	(50.85)	(19.647)	(34.785)
	[35.526]	[24.004]	[20.677]
n_menus	96.048***	52.873***	43.176***
	(14.988)	(5.791)	(10.253)
	[20.33]	[10.431]	[11.656]
treated	31.349	5.628	25.721
	(14.778)	(5.71)	(10.109)
	[25.156]	[7.693]	[18.486]
rank_true	-6.514	1.189	-7.703***
	(2.88)	(1.113)	(1.97)
	[4.613]	[2.511]	[2.58]
treatedwithin * rank_manipulation	32.313***	5.979***	26.334***
	(2.227)	(0.861)	(1.524)
	[5.275]	[1.29]	[4.123]
promoted_to_line	367.434***	104.716***	262.718***
	(27.635)	(10.678)	(18.904)
	[93.576]	[30.849]	[64.49]
demoted_from_line	-264.445***	-71.819**	-192.626***
	(30.13)	(11.641)	(20.611)
	[76.292]	[27.873]	[50.342]
$WeekDummies$	Yes	Yes	Yes
N	817	817	817
R-Squared	0.554	0.472	0.562
R-Squared Adj	0.534	0.454	0.542
F-Stat (p-value)	0	0	0

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in []
First Differences Estimator

ulating the rank, promoting and demoting movies to and from the new menu affect sales as before. As expected, trimmed is not statistically significant, as well as treated.

Table 7: Results eliminating sequences of treatments within the same cycle.

Subscribers	All	Standard	Premium
Model	FD	FD	FD
Variables	$leases_{it}$	$leases_{it}$	$leases_{it}$
(Intercept)	-5.68*	-2.752*	-2.928*
	(2.683)	(1.418)	(1.738)
	[3.027]	[1.663]	[1.729]
log(movie_age)	-12.366**	-11.933***	-0.433
	(6.003)	(3.172)	(3.889)
	[5.909]	[3.79]	[2.993]
n_menus	11.103***	5.499***	5.605***
	(1.819)	(0.961)	(1.179)
	[3.187]	[1.677]	[1.771]
treated	2.692	1.078	1.615
	(1.981)	(1.047)	(1.284)
	[3.29]	[1.532]	[2.482]
rank_true	-0.708	0.141	-0.849**
	(0.355)	(0.188)	(0.23)
	[0.843]	[0.621]	[0.352]
treatedwithin * rank_manipulation	2.914***	0.365	2.549***
	(0.323)	(0.171)	(0.21)
	[0.581]	[0.271]	[0.429]
promoted_to_line	39.786***	10.188***	29.598***
	(3.721)	(1.966)	(2.411)
	[6.676]	[2.55]	[4.855]
demoted_from_line	-31.377***	-6.851	-24.526***
	(4.719)	(2.494)	(3.058)
	[8.177]	[4.326]	[4.998]
trimmed	-0.405	0.184	-0.589
	(4.856)	(2.566)	(3.146)
	[4.934]	[2.371]	[3.712]
Week Dummies	Yes	Yes	Yes
N	762	762	762
R-Squared	0.456	0.265	0.49
R-Squared Adj	0.437	0.254	0.47
F-Stat (p-value)	0	0	0

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in [ ]
First Differences Estimator