Advertising and Word-of-Mouth in Motion Picture Industry*

Zakaria Babutsidze[1,2]

[1] SKEMA Business School, 60 rue Dostoievski, 06902 Sophia Antipolis, France
[2] OFCE Sciences Po, 250 rue Albert Einstein, 06560 Valbonne, France
zakaria.babutsidze@sciences-po.fr

Abstract. Motion picture industry is characterized by extensive advertising and word-of-mouth among (potential) consumers. We develop a simple computational model of consumer behavior to study the interaction between these two forces. WOM that propagates through fixed social network is affected by the mismatch between consumer's expectations and realized quality of the film. As a result, intensive advertising is running a risk of generating overly negative WOM, while moderate levels of advertising generate positive WOM that is complementary to advertising efforts. The most striking finding is that marginal returns to advertising can become negative high levels of advertising. This effectively means that for intensive advertising campaigns an additional commercial might result in the reduction of the audience.

1 Introduction

The motion picture industry has always advertised heavily [5]. But in recent years movie advertising budgets have been increasing fast. They have grown 50% between years 1999 and 2005 and hit 60% of the total production costs [1].

At the same time the industry is characterized by intensive word-of-mouth communication among (potential) consumers. This makes sharing of experience and influencing late-comers' decisions an easy task. Word-of-mouth is a powerful force which, through multiple exchanges can reach and influence large portion of the society [11, 4]. Word-of-mouth augments advertising in diffusion of information across the potential customers. However, unlike advertising, which has a clear aim of inducing people to buy a product, word-of-mouth does not have an ultimate target as it is un-coordinated collective effort. The important implication of this difference is that it is possible, and in fact quite plausible, that negative sentiment about the product diffuses through peer-to-peer interaction. Due to the fact that consumer interaction plays somewhat similar but at the

* Part of this work was done when I was at MERIT, Maastricht University and at the Center for Complexity in Business, University of Maryland. I thank both institutions for their hospitality and support. Discussions with PJ Lamberson, Bill Rand and Bulat Sanditov were particularly beneficial. I am also grateful for comments from the participants of various meetings in Maastricht, New York and Sophia Antipolis.

same time very different role from advertising, makes studying the interaction between these two forces interesting.

It is believed that negative word-of-mouth has effects of substantially larger size than its positive counterpart [13, 14]. This belief has been recently verified in a range of industries like airlines [12], online bookstores [6] and computer games [17]. Detrimental effects of negative word-of-mouth have been reported by early investigators like [15] and [16], who find that negative word-of-mouth significantly reduces perceived credibility of advertising as well as brand attitudes and purchase intensions. Dissatisfied consumer outrage has taken a central stage in deeper investigations such as [3]. Consumer efforts in response to dissatisfaction have been found to be even higher in case of business-to-business customers where each buyer is considerably larger in size [8].

Intuitively, this asymmetry across the sentiment directions can be understood by acknowledging the important role of time in purchase dynamics. Not all the people (willing to buy a product) buy a product immediately. Purchases are postponed and sales are stretched in time. Positive word-of-mouth can accelerate this process but only to a limit. However, due to the disappointment aversion [9], negative word of mouth has a power of permanently halting product sales at any point in time.

One other noteworthy characteristics of the motion picture industry is that it supplies an experience product. Therefore, a judgement about the quality of the product consumer is about to purchase can be imprecise. As potential consumers are aware of this fact, their perceptions are susceptible to word-of-mouth coming from consumers that have already seen a movie.

In this paper we present an agent based model in order to analyze the interaction between advertising and information diffusion through social networks. We combine above-listed features of the movie industry with the well-known psychological finding about the effects of disappointment and propose a simple model outlined in the following section.

2 The model

The economy consists of constant number (I) of consumers (indexed by i). A new film appears on the market. It has a (subjective) quality that is distributed across the population as $x_i \sim \mathcal{N}(\mu_x; \sigma_x^2)$ (or in general with an arbitrary distribution $x_i \sim PDF_x$). However, this quality is not known to consumers prior to seeing the movie. As we measure the quality from the standpoint of general public, rather than from the standpoint of a film critic, spurious relation between the quality and movie returns are removed in our framework [10].

Each consumer has an internal quality requirement y_i for going to the movies. She only goes to see the movie if her expectation for movie's quality is no less then y_i. This simple mechanism ensures that consumer behavior is consistent with a disappointment aversion as formulated by [9]. This variable is distributed across population as $y_i \sim \mathcal{N}(\mu_y; \sigma_y^2)$ (or in general case as $y_i \sim PDF_y$). This

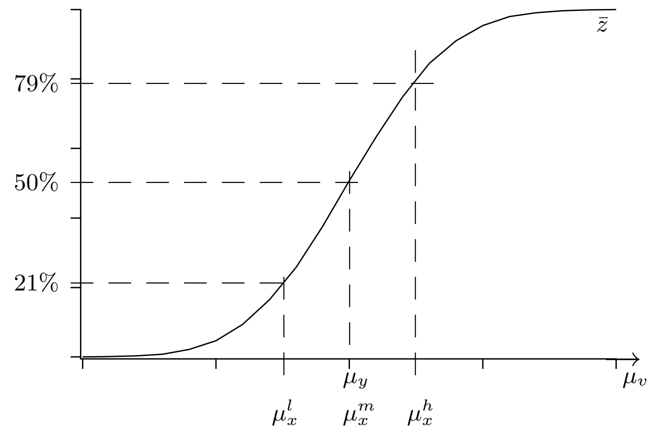


Fig. 1. Benchmark diffusion quantities. The plot is calibrated by the relationships $\sigma_x = \sigma_y = \sigma_v$ and $\mu_x^l - \mu_x^b = \mu_x^h - \mu_x^m = 1.25\sigma_x$

implies that consumers are heterogenous with respect to the disappointment aversion rate.

v_i^t is the belief that consumer i holds about the quality of a movie at time t. Initial beliefs about the quality of a movie are distributed in population as $v_i^0 \sim \mathcal{N}(\mu_v; \sigma_v^2)$ (or in general case as $v_i^0 \sim PDF_v^0$). Changes in beliefs as time progresses are incorporated into $v_i^t \sim PDF_v^t$.

If there is no social interaction or advertising, we can calculate how widely the given movie will be watched by the society. We can compute this value using a random matching mechanism

$$\bar{z} = \int PDF_v^0(z) \times CDF_y(z)\mathrm{d}z \tag{1}$$

$$= \frac{1}{4} \int \frac{1}{\pi\sigma_v^2} \exp\left(-\frac{(z-\mu_v)^2}{2\sigma_v^2}\right)\left(1 + \mathrm{erf}\left(\frac{z-\mu_y}{\sqrt{2\sigma_y^2}}\right)\right)\mathrm{d}z \tag{2}$$

Equation (1) and its special case (for v_i^0 and y_i being normally distributed) (2) give the share of consumer base that will watch a movie given the average initial belief that consumers hold about the quality of the movie. Equation (2) is plotted on figure 1 for the case when $\sigma_v = \sigma_y$.[3] Notice that due to the fact that consumers do not interact and exchange their impressions about the movie, the actual quality of the movie does not affect the number of people that will see it. On figure 1 we also identify the average accepted quality of the society – μ_y, which is measured on abscissa together with the initial average expected quality, μ_v.

[3] Throughout the whole paper we consider the arrangement when $\sigma_v = \sigma_y = \sigma_x$.

For the analysis in section 3 we use three types of movies. We define an average quality movie, a movie for which $\mu_x = \mu_y$. Which intuitively means that in absence of interaction the movie will be watched by half of the population. We denote the quality of an average (medium quality) movie by μ_x^m. Corresponding values of low and high quality movies are denoted by μ_x^l and μ_x^h, respectively. We define these products such that $\mu_y - \mu_x^l = 1.25\sigma_y$ and $\mu_y - \mu_x^h = -1.25\sigma_y$. In our setup this means that a low quality movie in the absence of consumer interaction will be watched by around 21% of the population, while a high quality movie will be seen by 79% of the potential customers. We also identify these three types of movies on figure 1.

Now we introduce two central forces in our morel that can affect consumer decisions – interaction and advertising. Let's take them one by one in a reverse order. Producers of the movie can advertise a product. Advertising is costly. Producers can advertise only before the movie hits the theaters. This is in line with the common practice of the industry – [7] find that 90% of advertising in motion picture industry is done before the release. They also find that pre-release advertising has a significant positive effect on expectations potential consumers hold about the movie. We assume that advertising is effective and it can increase μ_v. However, we assume that advertising cannot make tastes more homogenous, thus it cannot affect σ_v. Of course, driving μ_v to a higher level requires more spending. From the figure 1 one can clearly see that higher advertising expenditures would result into higher sales. However, this is only true when there is no communication among consumers.

To model the word-of-mouth interaction, we assume that there is a static, given social network that specifies the interaction structure among consumers. Information about the movie streams through this social network and affects the nodes (consumers). However, not all the nodes are functional at any time period. We assume that only the people who have not seen the movie update their beliefs. Once the person has seen the movie she has no reason to act on the information communicated to her by the social network. Therefore, the node corresponding to this consumer becomes dysfunctional – no new information passes through it.

Consider people going to the movie theatre one by one (i.e. each time period only one person can go to see the movie). Who goes at time t is randomly selected from the people whose $v_i^t \geq y_i$. Once a person goes to a movie, she realizes x_i, and deduces the final impression about the movie $v_i^T = x_i + a(x_i - v_i^t)$, where parameter $a \geq 0$ controls the strength of the effect of the mismatch between the expectations and subjective quality on the final impression of a person.

Once a person exits the theatre, she communicates v_i^T to her friends, who update their beliefs according to $v_j^t = v_j^{t-1} + b(v_i^T - v_j^{t-1})$, where $b \in [0;1]$ is a measure of how much people trust the judgement of their social contacts. These friends communicate v_j^t's further to their contacts. Further down the line people update their beliefs with $v_m^t = v_m^{t-1} + b^k(v_n^t - v_m^{t-1})$, where m receives the information from n, and k is the shortest (currently functional) path length from the person that went to the movies (i) to m. Modeling consumer interaction this way implies that social distance affects negatively the weight that consumers put on each other's judgements. After the information diffusion the node corresponding to the consumer i becomes dysfunctional. As a consequence functional social network changes.

3 Results

In order to discuss the implications of the model we run the economy until $v_i^t < y_i$ for all the people who have not seen the movie and measure the success of the setup by the share of consumer population that has seen the movie.

We employ the Monte-Carlo methodology. We run each setup 200 times for different random initial values and average them. Therefore, any point on the figures presented below comprises the average of 200 runs. Standard deviations in all the cases are extremely small, therefore we do not present them in plots.

In what follows we concentrate on the effects of the level of trust in the society and the topology of social network. We also check the robustness of our results by changing other parameters if the model. Important insights from these exercises are reported as additional results in the paper.

Although our variables, except the network architecture, are continuous, for the purpose of the presentation of results we discuss only three values for each parameter - low, medium and high. Therefore we define three levels of trust – b^l, b^m and b^h; three values of mismatch effect – a^l, a^m and a^h; and three values of network density – d^l, d^m and d^h. We will explore three types of social network architectures: lattice (that we denote by L), preferential attachment (denoted by P) and random network (denoted by R).

Every run has five parameters that specifies the characteristics of the economy. These are the quality of the movie, the trust in the society (strength of WOM), the strength of the mismatch between expected and realized quality of the movie, social network architecture and its density. Therefore each setup can be characterized by a set of five values $(\mu_x; b; a; A; d)$, where $A \in \{L, P, R\}$ denotes the topology of the social network. This notation will be used in the process of reporting the results.

We perform numerical simulations in the same setup as we have plotted the figure 1. Which are $\sigma_x = \sigma_y = \sigma_v$ and $\mu_x^m - \mu_x^l = \mu_x^h - \mu_x^m = 1.25\sigma_x$. Exact numerical values for the parameters do not affect the results. But we anyway report them in table 1.

3.1 The effect of trust

To demonstrate the effects of the trust in the society we compare the results of numerical simulations to the benchmark diffusion quantities given in figure 1. This is due to the fact that no communication can be viewed as communication with zero trust the difference between the two quantities is the measure of the effect of trust.

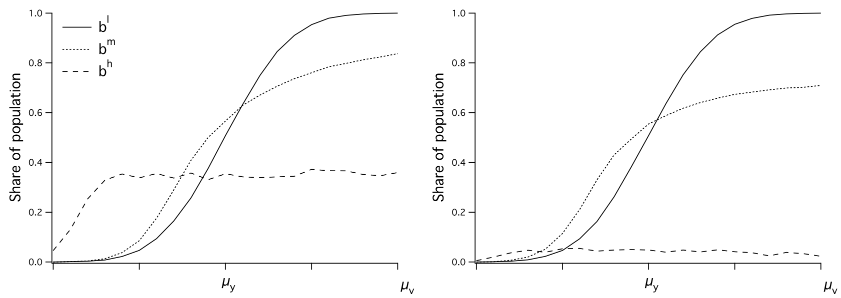
parameter	description	value
μ_y	The average accepted quality	20
μ_x^l	The average quality of a poor movie	18.75
μ_x^m	The average quality of a moderate movie	20
μ_x^h	The average quality of a good movie	21.25
σ_y	The variance of the accepted quality	1
$\sigma_x^l = \sigma_x^m = \sigma_x^h$	The variance of the quality	1
σ_v	The variance of initial quality expectations	1
a^l	Low effect of mismatch	0
a^m	Medium-size effect of mismatch	0.5
a^h	High effect of mismatch	1
b^l	Low level of trust	0
b^m	Medium level of trust	0.5
b^h	High level of trust	1
d^l	Low density (measured by average degree)	4
d^m	Medium density (measured by average degree)	20
d^h	High density (measured by average degree)	60
I	The number of consumers	1 000

Table 1. Parameter values for the numerical analysis.

Figure 2 presents the total sales of an average quality movie in case of communication taking place on lattice architecture. The left panel presents the situation when social network is sparse, while the right panel presents the situation in a denser network.[4] The sales in the benchmark situation are given on both of the panels. Recall that for this movie $\mu_x = \mu_y$.

As one can clearly see from figure 2 the communication moderates the effects of advertising. For low levels of advertising advertising, when μ_v^0 is low (compared to the actual quality of the movie) communication complements the advertising and helps increase the sales. However, if advertising is too fierce, communication decreases the sales. This is intuitive as fierce advertising drives consumer expectations up and a typical viewer gets disappointed with the movie. As a consequence negative word-of-mouth spreads and decreases the likelihood of other people seeing a movie. The higher the level of trust between members of the society (b) the more pronounced is the communication effect.

[4] In this an every subsequent figure the industry characteristics presented in [square brackets] is the one that is different in setups on left and right panels. For example in this case (the case of figure 2) density of the network is presented in blue. Which means that left and right panels of the figure are produced in the same setup except the density of the social network.



$(\mu_x^m; \bullet; a^l; L; [d^l])$ $(\mu_x^m; \bullet; a^l; L; [d^m])$

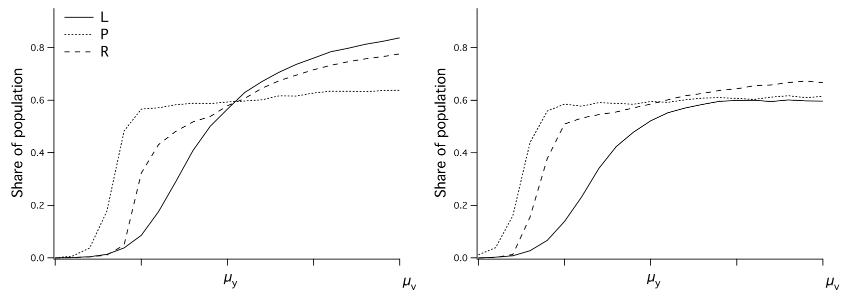
Fig. 2. The effect of trust on total sales.

Besides, the difference between two panels demonstrates the influence of the network density on effects of trust. Recall that after every act of consumption one node in the social network becomes dysfunctional. This means that at some point in time parts of a social network might get disconnected. In this environment higher density means that for any given number of viewers larger part of the network stays connected. If network is disconnected in two subnetworks the information coming from any given viewer cannot reach viewers in a subnetwork disjoint to the one this consumer belongs to. Then, as sparse networks get easily fragmented, they localize the information at later stages of industry development. This is the reason why sales stay at a low level for any level of advertising in setups with denser networks and high level of trust - WOM from dissatisfied customers can reach large parts of the network.

Figure 2 depicts the scenarios when the mismatch between anticipated and actual quality does not affect the final impression of the viewer ($a = 0$ and therefore $v_i^T = x_i$). When the mismatch affects the sentiment that viewer is diffusing through social network, variance in word-of-mouth increases. This means that negative word-of-mouth becomes even more negative, while positive word of mouth becomes even more positive. As it can be anticipated, due to the polarization of the sentiments helping hand of WOM towards the advertising at low levels increases (because most of the word-of-mouth is positive for advertising at low intensities). But for higher quantities of advertising WOM becomes more effective deterrent of the sales. This is especially pronounced for higher levels of trust, and is present in networks of all densities [AR].[5]

The present model implies that higher quality movies can capitalize on overly positive WOM at low levels of advertising. This effect is amplified by the sparse social networks, that localize few disappointed viewers [AR]. When considering preferential attachment and random network topologies it turns out that in case

[5] Results marked with [AR] are important enough to be included in the paper. However, the paper format does not allow for presentation of the documentation of these results. Plots and other documentations for the results marked with [AR] can be obtained upon request from the author.



$\left(\mu_x^m; b^m; a^l; \bullet; [d^l]\right)$ $\left(\mu_x^m; b^m; a^l; \bullet; [d^h]\right)$

Fig. 3. The effect of social network topology on total sales.

of both of these networks advertisers can achieve higher returns for low levels of advertising and lower returns for higher levels of advertising (compared to the lattice) [AR]. The reason is that both of these networks are less structured compared to lattice and both of them have lower average shortest path length.

3.2 The effect of network architecture

In this section we study the effect of the social network topology. We have used three topologies: lattice, preferential attachment and random. Figure 3 presents the results for the constant values of a and b. For generating the left panel we have used a sparse network, for generating the right panel dense network has been used.

Figure 3 demonstrates the main implication of the model. For low levels of advertising preferential attachment network has highest returns out of all three topologies. This is due to the fact that the preferential attachment has smallest average shortest path length out of all three architectures. As WOM for low levels of advertising is positive - lower average shortest path length guarantees faster diffusion of positive sentiment through social network. However, for the high intensity of advertising the ranking of structures depends on the density of the network. For sparse networks lattice is the most beneficial structure, while for dense networks random network results in highest returns.

To understand why this is the case we have to notice the difference between high and low levels of advertising. Besides the fact that at low levels WOM is positive and in high levels it is negative there is another crucial difference. It is that high levels of advertising results in higher sales. Every time a sale takes place the node in the social network becomes dysfunct - it does not pass any further WOM. Therefore, higher sales result higher number of dysfunctional nodes. As sales are random in our model, each visit to the movie theatre can be viewed as a random error in the social network (*a la* [1]). At low advertising intensity these errors do not change the functional social network architecture much. However, at high sales functional network changes significantly in the process.

When networks are sparse clustering is low, which means there is not much redundancy in the network. Therefore, lattice networks easily get fragmented into disconnected sub-networks that localize the negative WOM and result in higher sales for higher levels of advertising. However, if the network density is high lattices are more resistant to the random attacks (due to high redundancy) and it is random networks that have higher likelihood of being fractured into disconnected sub-networks. Then, returns to advertising are the highest in case of dense random social networks at higher levels of advertising. In any case, networks generated by preferential attachment are more resistant to random errors than Erdos-Reny random graphs in line with the findings by [1].

These results carry over for different values of other parameters, and the difference becomes dramatic for high values of trust in the society [AR]. The reason for this is that trust amplifies the role of the shortest path length in the dynamics. If paths are long, low trust makes WOM decay quickly. However, if trust is high information sent by a node will reach (and influence) other nodes even if the path is long (except if it is infinite, which is the case if sender and receiver are in different disconnected components of the network).

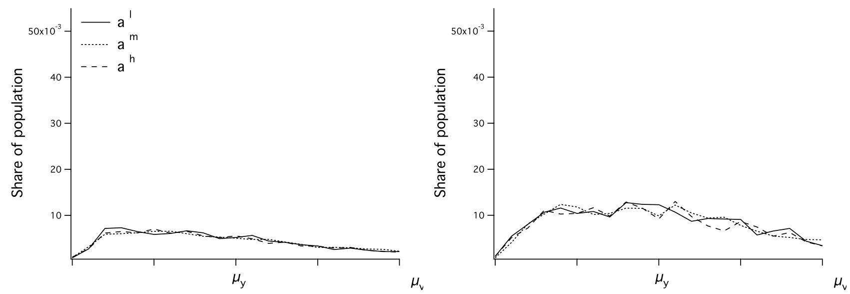
4 Returns to advertising

Heavy advertising practices in the movie industry raise the question of the optimality of these expenditures. As we can see from figures in this paper (e.g. the left panel of figure 2) movie sales level off after some point. Which means that additional advertising efforts do not increase sales. Therefore, marginal returns to advertising fall to zero. As marginal cost of advertising will never go to zero, we can claim that advertising with the intensity where marginal returns to product promotion are zero will be not optimal. It will be a waste of resources. Although we cannot pinpoint the optimal advertising level (as we do not have advertising costs in our model), we can be sure that the optimal rate of advertising in any arrangement has to be in the area where sales are still increasing.

However, our model has another interesting implication. Even if we would have assumed that marginal cost of advertising could fall to zero, the model implies that advertising efforts will still be bounded from above. Figure 4 demonstrates the finding that after certain level marginal returns to advertising become negative (i.e. more advertising results into less sales). This is due to the fact that too aggressive of an advertising campaign is bound to leave behind large numbers of disappointed people. And the negative word-of-mouth will discourage latecomers from going to the cinema.

The point after which the advertising becomes consumption-deterrent is reached at lower advertising levels for poor quality movies, but even hits can suffer from it. The tipping point lies at lower advertising levels for the denser networks and for the societies with high levels of trust [AR].

Finding is robust in changes of network topologies. However, one peculiarity emerges. Random social network in presence of low levels of a seems to maintain positive returns for much higher levels of advertising compared to the other two



$\left([\mu_x^l]; b^h; \bullet; L; d^h\right)$ $\left([\mu_x^m]; b^h; \bullet; L; d^h\right)$

Fig. 4. Total sales for different levels of advertising.

architectures (lattice and preferential attachment), which behave similar to each other and to random network with high values of a [AR]. The reason for this is that for low expectation mismatch effect (a) WOM is not that negative therefore consumption continues at up to a certain level. However, because networks are dense, out of all the architectures random network will suffer the most. Which means that average shortest path length will increase in random networks. And negative WOM once again will get localized that would allow certain parts of the network to continue consuming even further. However, when a is high, WOM is overly negative and consumption stops at lower level, earlier than random network gets fractured into pieces that can localize WOM. Similar behavior does not emerge in other two topologies. This is due to the fact that scale free networks (that are generated by the preferential attachment algorithm) are known to be good resistants to random errors [1]. In our model consumption can be viewed as the error in network. Lattices are also good resistants because they posses a structure that allows them to avoid fracturing into pieces.

As we can see in our model aggressive advertising might become consumption-deterrent due to the exceedingly negative WOM that it generates. [2] provides a different reason for consumption-deterrent advertising. [2] studies the information provision role of the advertising on proliferated markets. Unlike our setup, in that model advertising is not content free. It provides information about the characteristics of the product. Therefore, in the presence of multiple options advertising increases the likelihood of better match between a consumer and a product. As a consequence, (non-content-free) advertising might reduce sales if consumers realize that the product is not what they would like to buy. So, in the setup of [2] advertising can be consumption-detering due to the mismatch between the consumers preferences and (at least) "partially-observable" characteristics of the product, while in our setup advertising can be consumption-deterring due to the mismatch between the information received through advertising and through word-of-mouth.

5 Conclusion

In this paper we have presented a simple computational model of motion picture viewer behavior in order to analyze the interaction between advertising and word-of-mouth that diffuses through social interactions. The most striking result is that if advertising is too intensive, marginal returns to advertising becomes negative. This effectively means that advertising results in smaller number of consumers going to see the movie. This is due to the fact that advertising drives viewer expectations high and as a consequence leaves large number of them disappointed due to the mismatch between their expectations and realized movie quality. This triggers negative WOM that affects consumption intensions of following (potential) customers.

Empirical research into the motion picture industry has found that marginal returns to advertising is indeed too low, however not negative. [7] find that on average movie sales increase only by 0.65 dollars for an additional dollar spent on advertising. Looking at these results from the lense of the model discussed in this paper, we can argue that advertising is at inefficiently high levels.

An important shortcoming to the present model is that although it tells us that firms over-advertise it offers no explanation of why this can be the case. The reason for this is that the model is stylized and discusses returns to advertising in absence of competitors. Advertising in our framework works only for creation of the market. However, in the real world advertising has is also used as a competition tool. Therefore decisions about the intensity of advertising become strategic. In these environments it is easy to imagine that advertising is growing at inefficiently high levels due to producers trying to keep up with the competition for market shares. However, our research warns that these kind of overly aggressive advertising campaigns can reduce not only the market share of a movie, but also the size of the whole movie market.

References

1. Albert, R., Jeong, H., Barabasi, A.-L.: Error and Attack Tollerance of Complex Networks. Nature. **406** (2000) 378–382
2. Anand, B., Shachar, R.: Advertising, the Matchmaker. RAND Journal of Economics. (forthcoming)
3. Bechwati, N. N., Morrin, M.: Outraged Consumers: Getting Even at the Expense of Getting a Good Deal. Journal of Consumer Psychology. **13** (2003) 440–453
4. Brown, J., Broderick, A. J., Lee, N.: Word of mouth communication within online communities: Conceptualizing the online social network. Journal of Interactive Marketing. **21** (2007) 2–20
5. Caves, R. E.: Creative Industries: Contracts between Arts and Commerce. (2001)
6. Chevalier, J., Mayzlin, D.: The effect of word of mouth on sales: online book reviews. Journal of Marketing Research. **43** (2006) 345–354
7. Elberse, A., Anand, B.: The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market. Information Economics and Policy. **19** (2007) 319–343

8. Ferguson, J. L., Johnston, W. J.: Customer response to dissatisfaction: A synthesis of literature and conceptual framework. Industrial Marketing Management. **In Press** (2010)
9. Gul, F.: A Theory of Disappointment Aversion. Econometrica. **59** (1991) 667-86
10. Holbrook, M. B.: Popular appeal versus expert judgements of motion pictures. Journal of Consumer Research. **26** (1999) 144–155
11. Lau, G. T., Ng, S.: Individual and Situational Factors Influencing Negative Word-of-Mouth Behaviour. Canadian Journal of Administrative Science. **18** (2001) 63–178
12. Lou, X.: Consumer Negative Voice and Firm Idiosyncratic Stock Returns. Journal of Marketing. **71** (2007) 75–88
13. Mahajan, V., Muller, E., Kerin, R.: Introduction Strategy for New Products with Positive and Negative Word-of-mouth. Management Science. **30** (1984) 1389–1404
14. Park, C., Lee, T. M.: Information direction, website reputation and eWOM effect: A moderating role of product type. Journal of Business Research. **62** (2009) 61–67
15. Singh, J.: Voice, Exit, and Negative Word-of-Mouth Behaviors: An Investigation Across Three Service Categories. Journal of the Academy of Marketing Science. **18** (1990) 1–15
16. Smith, R. E., Vogt, C. A.: The Effects of Integrating Advertising and Negative Word-of-Mouth Communications on Message Processing and Response. Journal of Consumer Psychology. **4** (1995) 133–151
17. Yang, J., Mai, E. S.: Experiential goods with network externalities effects: An empirical study of online rating system. Journal of Business Research. **63** (2010) 1050–1057