

The effect of social interactions in the primary consumption life cycle of motion pictures

César A Hidalgo R^{1,2}, Alejandra Castro³
and Carlos Rodriguez-Sickert^{4,5}

¹ Department of Physics and Center for Complex Network Research,
University of Notre Dame, Notre Dame, IN 46556, USA

² Kellogg Institute, 130 Hesburgh Center, Notre Dame, IN 46556, USA

³ Department of Physics, University of Michigan, Ann Arbor, MI 48104, USA

⁴ Department of Sociology, Pontificia Universidad Católica de Chile,
Vicuña Mackenna 4860, Macul, Santiago, Chile

⁵ Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

E-mail: crodrigs@uc.cl

New Journal of Physics **8** (2006) 52

Received 7 December 2005

Published 10 April 2006

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/8/4/052

Abstract. We develop a ‘basic principles’ model which accounts for the primary life cycle consumption of films as a social coordination problem in which information transmission is governed by word of mouth. We fit the analytical solution of such a model to aggregated consumption data from the film industry and derive a quantitative estimator of its quality based on the structure of the life cycle.

Contents

1. Introduction	2
2. The model	3
2.1. An atomized society with no supply restrictions	4
2.2. Cultural consumption and the effect of social interactions	4
3. Empirical validation	7
4. Conclusion	10
Acknowledgments	11
References	11

1. Introduction

The study of the consumption of cultural goods in general, and that of films in particular, has been traditionally restricted to total demand empirical studies [1]–[3]. In these studies variations in the quality of the goods have been consigned to a residual status, focusing in the effect of prices and income as the main explanatory factors. More recent studies have attempted to incorporate quality obtained from individual expected values [4] or, alternatively, as a macro variable obtained ex-post from critics' review indexes [6] or from online reviews [4]. However useful, demand studies such as the ones surveyed above do not account for the dynamics of the consumption life cycle of the cultural goods and, more importantly, they do not discuss the ways in which the structure of this cycle is affected by the transmission of information from members who have had access to the cultural good to potential consumers.

Within the economic literature, the focus on social interaction and the dynamics of consumption of cultural goods has been mainly confined to 'fashion cycles' [8, 9].⁶ In our case, however, the demand for a product does not decrease due to the consumption of it by other agents. The movie does not become worn out, like a fashion design. Outside the economic field, the sociologist Lipovetsky [7] has focused his work on the macro structure of the life cycle, mainly in terms of its duration. Although the argument is not formalized, it is in concordance with the way in which novel technologies diffuse [10]–[12]. Word of mouth, in the context of the adoption of certain technologies, can be justified as a valid mechanism of social influence. Given the scale economy that characterizes certain technologies, e.g., database software, it is impractical (or too expensive) to do small-scale experiments and, thus, word of mouth becomes an important part of the diffusion process of these investment goods [13].

The working hypothesis of our approach is that potential film consumers face a similar problem to agents who decide whether or not to adopt a novel technology. However, certain features of the former problem give raise to different empirical regularities at the macro level. 'Early adopters' within the film context do not face the risks normally associated with the adoption of a novel technology and thus an important fraction of them, solely based on the pre-opening expectations generated by the media and advertising, will be willing to consume the product before peer opinions become available (i.e., before the novel technology is tested). Thus, in contrast to the standard S-shaped behaviour associated with the adoption of a novel technology, films tend to have a relatively high initial attendance (as a fraction of total attendance) which usually decreases all along the consumption cycle. Only unexpectedly successful small productions or independent films show a consumption life cycle that resembles that of the adoption of a novel technology. In such cases, the adventurous first attendants play a similar role of early adopters of a successful novel technology.

In this paper, we develop a 'basic principles' model which accounts for the primary life cycle consumption of films as a social coordination problem in which information transmission is governed by word of mouth. Rather than positing an *ad hoc* Bass-type diffusion account of the aggregated dynamics of the consumption life cycle as in [14], this dynamics is analytically derived from an explicitly defined model of social interaction. The resulting dynamics account for both types of cycles: S-shaped cycles and continuously decreasing cycles. The model is empirically validated using data from the film industry. The behaviourally based approach followed here allows us to obtain estimates of the transmitted average opinion about the quality of the movies

⁶ A notable exception is [5]. However, rather than focusing on the consumption life cycle itself, they focus on the way in which supply and demand interact.

as perceived by their targeted audience.⁷ Thus, in contrast to total demand studies in which quality is an exogenous variable, here it is obtained from the structure of aggregated behaviour.

2. The model

We now present a model of cultural consumption based on the following assumptions.

- Each agent goes to see a movie at the cinema only once. (We neglect the probability of going twice assuming that the probability of going to the cinema to see the same movie n times is a rapidly decaying function of n .)
- The probability that an agent goes to the movies is affected by the interactions with agents who have already seen it.

The different quantities associated with the model will be expressed using the following notation. $N(t)$ will represent the number of agents that have not watched the movie at discrete time t , while $P(t)$ will represent the probability that an agent who has not watched the movie decides to attend it at discrete time t . The observable quantity is the number of agents who have seen the movie at discrete time t . We will denote this quantity $A(t)$. In our model, it will be given by the product $A(t) = P(t)N(t)$, which is nothing more than the expected number of agents attending the film.

Before the movie is available for its potential consumers, agents have a prior conception about its quality, which comes from the information of pre-observable features such as its budget, the cast, and advertisement. The availability of this information will depend on the marketing strategies of the producers and distributors. Irrespectively of the way in which these strategies are conceptualized, i.e., whether publicity is taken as an information provider or as a persuasive device, its effects on our model are equivalent, affecting the prior conception about the movie in question and, therefore, the agents' likelihood of actually attending it at the beginning of the process. We will denote this likelihood by P_0 , and we will call the initial target population N_0 , which represents the total number of potential attendants of the good before it becomes available.

Although the consumption life cycle of some cultural goods can be potentially unbounded (we still buy copies of *Don Quixote*), the aim of this work is to understand the dynamics of the primary life cycle, which finishes when per period consumption decreases below a certain threshold relative to its premiere level. A less ambiguous case is that of performing arts like theatre, where producers are forced to cancel presentations when box-office revenues goes below fixed costs. The alternative cost associated with the availability of new options, an effect which also applies to the film industry or best-seller book industry, strengthens the limited duration life cycle that characterizes aggregated consumption in these environments. Sedgwick [15] claims that more than 70% of film revenue is derived from non-theatrical resources. Indeed, our approach does not neglect the importance of subsequent phases of the consumption cycle. Given that the success (and length) of these subsequent phases are heavily dependent on the primary cycle, the latter is worth an investigation on its own.

⁷ The targeted audience is a somewhat elusive concept which involves well-defined niches (such as romantic comedies) together with the expectations generated by pre-opening sources of information about the quality of the movie.

2.1. An atomized society with no supply restrictions

The simplest case is that of an atomized society where agents' decisions are independent from each other, both in terms of information (the opinion of agents who have watched the movie do not influence the decision of potential consumers) and in terms of consumption (there are no network externalities associated with timely coordinated consumption). In this case, we will also assume that restrictions on the supply side do not apply. The whole targeted population could simultaneously go to the cinema and the horizon of its exhibition has no time limit.

In this case the probability of attending the cinema does not depend on time, and it is equal to its initial value P_0 . The system dynamics can be described by considering that the expected attendance at time t is given by the product between P_0 and $N(t)$; and that at the beginning of the process no agents have seen the movie. After the first time step some agents will have attended the movie, and we will therefore have to subtract them from the ones who have not. Thus the temporal variation of attendance will be given by

$$N(t + 1) = N(t) - P_0 N(t). \quad (1)$$

If we approximate the discrete variables by continuous ones and notice that in the first time step attendance is given by $N_0 P_0$, we can conclude that the general solution of equation (1) has the traditional form

$$N(t) = N_0 e^{-P_0 t}, \quad (2)$$

$$A(t) = N_0 P_0 e^{-P_0 t}. \quad (3)$$

In the next section we will see how social interactions alter this behaviour.

2.2. Cultural consumption and the effect of social interactions

Once the effect of social interaction is considered, $P(t)$ becomes a dynamical variable and its change in time is now due to two main contributions. The first one is associated with the transmission of information about the observed value of the movie, and it represents the change in the probability of attending the movie induced by the information about it transmitted between agents. After the first period of attendance, residual potential consumers have access to the opinion of the first period audience. This effect is cumulative as the residual consumers have potential access to the opinions of audiences from the previous time steps.⁸ We assume that the change induced in $P(t)$ by this effect will be proportional to the probability that agents who have not seen the movie meet with agents who have. The proportionality constant associated with this term will represent the strength of it. The second contribution is associated with coordinated consumption. In the case of performing arts like cinema, agents usually attend in groups; agents who have not seen the movie are less likely to go if they are not able to find other agents to keep them company. The change induced by this effect will always reduce attendance likelihood, and

⁸ In our model, word of mouth is a first order process (only opinions of effective attendants are paid attention to). Alternatively, one could consider an n th order process where opinions are transmitted to residual consumers also via agents who have not yet participated in the consumption of the cultural product. Qualitative properties of an n th order process in which the effect's strength diminishes over time, would be equivalent to the first order process investigated here. Wu *et al* [16] have shown that information flow remains bounded as long as the likelihood of transmitting information remains sufficiently small.

it is also proportional to the probability that agents who have attended the movie meet with ones who have not.⁹ Therefore, when considering the effect of social interaction, we write the change in the probability of consuming the product as

$$P(t+1) = P(t) + \sum_t (S_+(t) - S_-(t)) \frac{A(t)}{N_0} \frac{N(t)}{N_0}, \quad (4)$$

where S_+ represents the proportionality constant associated with the effects of information flow while S_- represents the social coordination effects. From now on we will focus on the case in which S_+ can take positive and negative values representing the fact that information transmitted between agents can stimulate or inhibit the future attendance of other agents, whereas S_- will always contribute negatively. This is because coordinated consumption can only reduce the likelihood of going to the cinema for un-coordinated agents. It is worth noticing that both terms act directly on the population, altering the likelihood that agents will attend or purchase the product. Although agents are not explicit rational optimizers who make their decision over the basis of a Bayesian update of their beliefs about the movie's quality,¹⁰ one could understand this model as the outcome of rational searching behaviour in an incomplete information environment.

A different approach for the diffusion of innovations is to consider site percolation on regular lattices [17, 18]. These models of social percolation are based on the assumption that agents occupy the vertex of a regular lattice in d dimensions. Innovations diffuse by percolating the lattice according to a certain quality value which, when above the percolation threshold, generates a massive consumption cluster. Both social percolation models and our model attempt to account for word of mouth. We assume homogeneity in the communication process between potential consumers, i.e. in our model everybody can speak to everybody. In contrast, communication in the social percolation model is spatially restricted to the neighbourhood of a lattice vertex. The effective mechanism of transmission is probably somewhere between the diffusion process configured in our model and social percolation. It is true that one speaks more to closely related agents, but, as the abundant literature in small-world networks [19, 20] suggests, word of mouth effects are not confined to a local neighbourhood. Which model is more appropriate might also depend on the type of movie analysed. One of the implications of the social percolation approach is that products which take-off (movies in our case) are those which happen to be planted in a cluster rich in potential buyers. Whether this may well be the case in independent productions in which the first people who watch the movie (and the cluster to which they belong to) might have a crucial influence in future aggregated consumption, it is less likely in studio production where initial attendance can easily exhaust the niches of a structured population.

Indeed, if more detailed information was available about the effective paths of communication between potential consumers a more comprehensive account of social structure could be incorporated using the framework provided by network theory. Social networks exhibit topological properties which consider a number of relevant features which a lattice cannot account

⁹ The latter contribution to change can also be associated with other social behaviours. An example of this is that agents do not consume coordinately just because of company, but can do so independently but at similar times, so they can discuss the product in question later.

¹⁰ Ellison and Fudenberg [26] develop a Bayesian model of learning in which agents update their beliefs about the quality of the product after meeting other agents. In their model, is the updated belief that affects the agents' optimization process.

for such as the scaling properties of its degree distribution [21, 22], the community structure [23, 24] and the short average path length [25] to name a few. In this paper, we present a model that does not take this substrate into account, but could be considered as a mean field solution of the structural model.

2.2.1. Analytical solution. The continuous version of the system can be solved analytically in the case in which S_+ , S_- do not depend on time and the effects are not cumulative. The latter assumption can be partially justified by arguing that agents are more likely to express their opinion just after attending a movie. Our main reason not to include cumulative effects is, however, to keep the model analytically tractable.¹¹ In this case (1) and (4) reduce to

$$N(t+1) = N(t) - P(t)N(t), \quad (5)$$

$$P(t+1) = P(t) + (S_+ - S_-) \frac{A(t) N(t)}{N_0 N_0}, \quad (6)$$

which in the continuous case can be represented by

$$\frac{dN}{dt} = -PN, \quad (7)$$

$$\frac{dP}{dt} = \sigma PN^2, \quad (8)$$

where

$$\sigma = \frac{S_+ - S_-}{N_0^2}.$$

To find a solution we notice that equation (7) can be substituted into equation (8) giving us a third differential equation relating N and P which can be solved and used to introduce the initial conditions of the system. This relation can be expressed as

$$P(t) = -\frac{\sigma}{2}(N(t)^2 - N_0^2) + P_0. \quad (9)$$

Replacing (9) back on equation (7) and integrating it, we find that the solution of the system is given by

$$N^2(t) = \frac{BN_0^2}{N_0^2 - (N_0^2 - B)e^{B\sigma t}}, \quad (10)$$

$$P(t) = -\frac{\sigma}{2} \left(\frac{BN_0^2}{N_0^2 - (N_0^2 - B)e^{B\sigma t}} - N_0^2 \right) + P_0, \quad (11)$$

$$A(t) = N(t)P(t) = N(t)^{1/2} \left(P_0 - \frac{\sigma}{2}(N(t) - N_0) \right), \quad (12)$$

¹¹ Numerical analysis shows that cumulative effects do accelerate the process of transmission, but do not change qualitative properties of the model, i.e., the introduction of these effects is equivalent to time re-scaling.

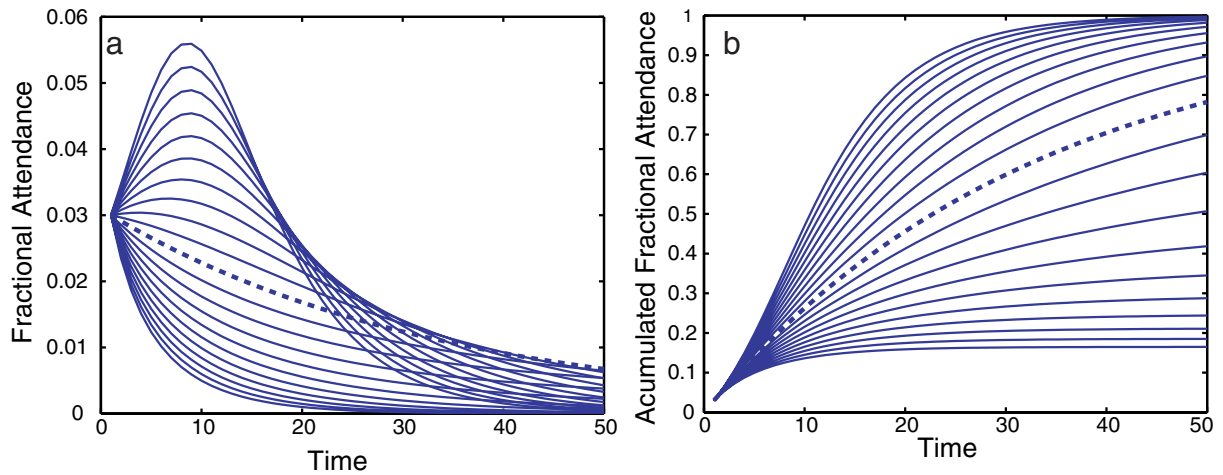


Figure 1. (a) Fraction of agents attending the cinema as a function of time according to the model described by equations (7) and (8). The segmented line represents the atomized behaviour described by equation (3) while the lines above and below represent the behaviour of positive and negative σ values respectively. (b) Shows the accumulated attendance, or in other words, the integral over time of (a).

where

$$B = N_0^2 + \frac{2P_0}{\sigma}.$$

Figure 1 shows the attendance as a function of time as well as the accumulated attendance normalized by the total target population ($A(t)/N_0$ and $\sum_t A(t)/N_0$). The segmented line represents the atomized behaviour occurring when $P(t)$ remains constant throughout the process. This line also divides the behaviour of the system in two: the solutions that lie above it are examples of systems which are dominated by a strongly favourable flow of information stimulating the consumption of the product; conversely, the lines that lie below represent systems that are dominated by coordination effects. In summary, we can see that two classes of behaviour are predicted. The first class emerges for negative σ values and is characterized by a monotonic decay with an exponential tail (notably, for this class of cycle, different values of σ lead to dramatic differences in the final number of spectators). When the strength of social interactions σ is large enough, a second class of behaviour emerges. In this case a period of increasing attendance exists until a maximum is reached. After this, the first class of behaviour develops. The latter case has an accumulated behaviour that resembles the standard ogive S-shape, which characterizes technological diffusion as accounted in Bass-type models.

3. Empirical validation

The model represented by equations (5) and (6) was validated by contrasting it with the US box-office data available on the Internet Movie Database (IMDb) Web site. We considered the

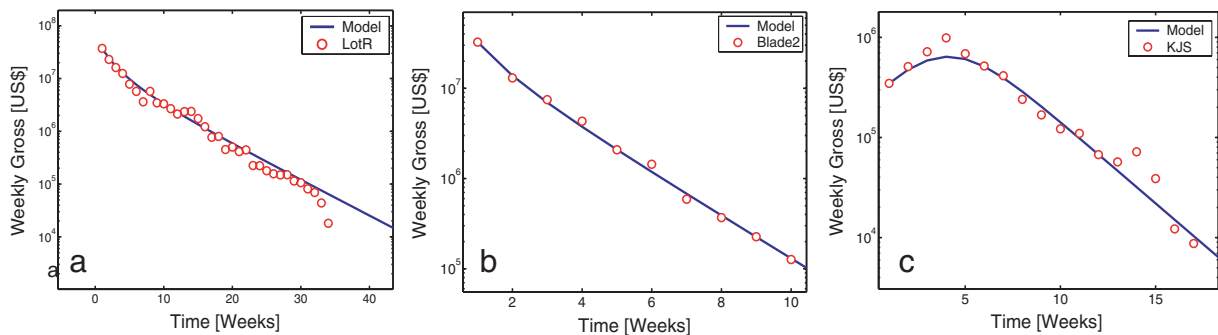


Figure 2. Three examples of the fitting procedure are shown for (a) *Lord of the Rings I* (b) *Blade II* and (c) *Kissing Jessica Stein*.

44 movies with the highest budgets of 2003 as our sample set and performed a χ^2 estimation procedure on all movies in order to find the parameter set which most accurately fitted the empirical results.

The model has three free parameters, the two initial conditions N_0 and P_0 , and the constant representing the strength of social interactions σ . The parameters used in the χ^2 minimization process were σ and N_0 while P_0 was determined by matching the first data point on the data set with the first data point in the model- $P_0 = (\text{First} : \text{Data} : \text{Point})/N_0$. This reduces the number of free parameters to two. The data set considered did not contain weekly attendance as a function of time but the weekly gross collected by each movie. We consider that the amount of money collected in the box office by a movie is a linear function of the number of people that attended it. This allows us to do our empirical analysis based on the weekly amount of money collected by a particular movie, which we claim to be equivalent to a similar analysis based on the number of agents attending it. Figure 2 shows three plots comparing the model with the empirical data. Both *The Lord of the Rings: Fellowship of the Ring* and *Blade II* represent examples of the first class of behaviour identified in the previous section. It is worth noticing that the slope followed by the data changes in time. The first points have a negative slope, which is considerably more pronounced than the ones in the tail. A simple exponential fit would be accurate only in one section of the curve, but would fail to fit both behaviours that were expected from the model and that appear to be present in real life.

In order to interpret the relative values of σ , and therefore compare different motion pictures, it is important to consider the behavioural class in which the particular films are located. Blockbuster movies have a high initial attendance, which causes population finite size and social coordination effects to be very strong. This implies that movies in this category will usually have a negative value for σ , and the magnitude of it will represent the observed value of the product. Small negative values are associated with movies in which the decay is not accelerated due to the observed quality, but exhibit some acceleration due to social coordination effects. On the other hand, large negative values map onto accelerated decays that in our model are due to social coordination effects plus poor values in the observed quality of the product.

In the second class of behaviour, the premier level of attendance is low, thus coordination effects do not act strongly on the system, because of the initially slow depleting of potential customers. In this case, small values of σ represent a movie in which the decay is not accelerated

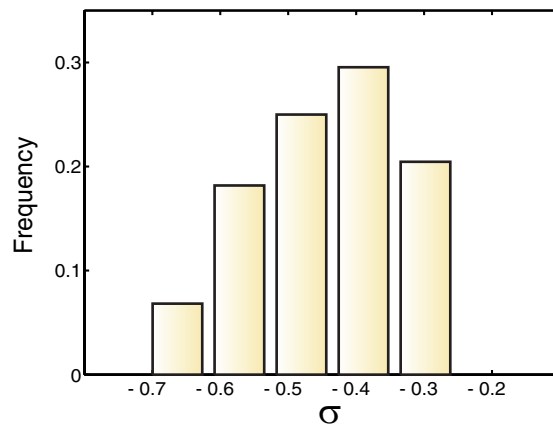


Figure 3. Distribution of the social influence parameter σ for the 44 studied movies.

or damped by social effects. On the other hand when relatively large values of σ are observed the attendance increases during the first time steps. This increase in consumption is due to the fact that the movie was well evaluated and that the pool of possible attendants was not depleted during the first couple of weeks.

The distribution of revenues has been empirically studied [5, 27] also by using the IMDb data set. In [27] it is shown that the opening gross as well as the total gross (scaled by their respective averages over all the movies released that year) show a power-law behaviour with the same exponent. The similarity of these two curves, they argue, is associated with a strong degree of correlation between the income of a movie at its opening, and its total income. The reported correlation shows that pre-opening public sources of information can determine the whole cycle of consumption. Based on this, they argue that movies can be classified into *blockbusters* having both high opening and high total gross and *bombs* (or *flops*). However, not only the existence of type I cycles—which they refer to as *sleepers*—but significant differences for the values of σ within type II cycles show that initial audience and total audience do not go always together. Indeed, it can be argued that when public sources of information (such as the cast or film reviews) do not accurately fulfil expectations or when this information is scarce, word of mouth has its greatest impact.

From the empirical point of view and after observing the whole data set it becomes apparent that the values of σ that we found tend to lie in a well-defined distribution. Figure 3 shows the distribution of values for σ obtained using this data set and the procedure explained above. A negative bias is seen when we observe the distribution of σ presented in figure 3. This negative bias tells us that in most cases, the coordination effect, which could also be associated with the underlying effect of competitions in the market for cultural goods, dominates the dynamical properties of the system. On the basis of the conjecture that coordinated consumption has the same effect on all motion pictures, we can interpret that the value extracted for σ represents the actual effect that the flow of information has on the film, and therefore on its consumption life cycle. This would indicate that coordinated consumption only introduces a shift in the value of σ and that the deviation from this well-defined mean represents the actual value of the movie as given by the targeted audience. For instance, the differences in the estimated parameter for *The Lord*

of *the Rings* ($\sigma = -0.3$) and *Blade II* ($\sigma = -0.6$), would indicate that the former film was much better received than the latter one. For *Kissing Jessica Stein* ($\sigma = 0.5$), a surprisingly enjoyable low budget romantic comedy, attendance evolved in the way a beneficial novel technology diffuses with the initial audience performing the role of early adopters. This reverse engineering definition of the observed value of subjective film reception can be used to refine the classification previously proposed. For us, *blockbusters* and *bombs* belong to type II cycles with average values of $\sigma = -0.45$. On the basis of significant deviations from this value, type II movies can be subclassified according to the way in which expectations are fulfilled/not fulfilled. Whereas extremely low values of σ can be referred to as a *disappointments* where expectations are not fulfilled (e.g., *Blade II*), σ values close to 0 can be referred to as *nice surprises* where expectations are more than fulfilled (e.g., *The Lord of the Rings*). What is interesting about this approach to film reception is that it is not based on the reported opinion of a film critic or an audience survey, but from the actual behaviour of agents.

4. Conclusion

A dynamical model representing the life cycle of motion pictures was introduced. The assumptions of the model were that

- agents do not go to the cinema to see a particular movie more than once, and that
- the probability that agents go to the cinema changes in a way which is proportional to the number of agents who have seen the movie times the ones who have not.

The first assumption gives rise to the exponential decay that characterize the tail of this process, while the second one allows the system to adjust its decay according to the social interactions present in cultural consumption. Consequently, it allows the model to accurately fit the different classes of observed behaviour, namely,

- a monotonic decay with an exponential tail, and
- an exponential adoption followed by an exponential tail which is traduced into an ogive S-shaped behaviour when the accumulated cinema attendance is observed.

In addition, under certain assumptions, our model can be used to infer a quantitative estimator of the subjective reception of a particular film. This estimate is directly inferred from the structure of the consumption life cycle.

Further research in this area can be pursued in a variety of ways. A natural extension of the model analysed here would be to consider the more general setting in which the agents face a number of cultural options and a limited budget for a given period of time. It is an area of empirical interest to investigate longitudinal processes in order to explore how the structure of the consumption life cycle has evolved along the last decades, and transversal processes to see the differences and correlations between the reception of particular films in different geographical regions.

Acknowledgments

We thank the comments of Samuel Bowles, Peter Richerson, Marcos Singer and the participants of the workshop on Evolution and the Social Sciences at the CEU, Budapest. We also acknowledge financial support from Fundacion Andes grant C-13960.

References

- [1] Baumol W and Bowen W 1965 *Am. Econ. Rev.* **55** 495
- [2] Moore T 1966 *Rev. Econ. Stat.* **48** 79
- [3] Blanco F and Baños-Pino B F 1997 *J. Cult. Econ.* **21** 57
- [4] Thorsby D 1994 *J. Econ. Literature* **32** 1
- [5] De Vany A and Walls W D 1996 *Econ. J.* **106** 1493
- [6] Urrutiager D 2002 *J. Cult. Econ.* **26** 185
- [7] Lipovetsky G 1987 *L'Empire de l'éphémère, la mode et son destin dans les sociétés modernes* (Paris: Gallimard)
- [8] Karni E and Schneider D 1990 *Am. Econ. Rev.* **80** 262
- [9] Pesendorfer W 1995 *Am. Econ. Rev.* **85** 771
- [10] Bass F M 1969 *Manag. Sci.* **15** 215
- [11] Mansfield E 1961 *Econometrica* **29** 741
- [12] Mansfield E 1968 *Industrial Research and Technological Innovation: An Econometric Analysis* (New York: Norton)
- [13] Moore G 2001 *Crossing the Chasm: Marketing In High-Tech Industries* (New York: Harper)
- [14] Jones M J and Ritz C J 1991 *Int. J. Res. Marketing* **8** 91
- [15] Sedgwick J 2002 *J. Econ. Hist.* **62** 676
- [16] Wu F, Huberman B A, Adamic L A and Tyler J R 2004 *Physica A* **337** 327
- [17] Stauffer D and Weibusch G 2003 *Int. J. Mod. Phys. B* **17** 5495
- [18] Goldenberg J, Libai B, Solomon S, Jan N and Stauffer D 2000 *Physica A* **284** 335
- [19] Milgram S 1967 *Psychol. Today* **2** 60
- [20] Watts D J, Dodds P S and Newman M E J 2002 *Science* **296** 1302
- [21] Barabasi A-L and Albert R 1999 *Science* **286** 509
- [22] Albert R and Barabasi A-L 2002 *Rev. Mod. Phys.* **74** 47
- [23] Newman M E J and Girvan M 2004 *Phys. Rev. E* **69** 026113
- [24] Newman M E J 2003 *Phys. Rev. E* **67** 26126
- [25] Watts D J and Strogatz S H 1998 *Nature* **393** 440
- [26] Ellison G and Fudenberg D 1995 *Q. J. Econ.* **110** 93
- [27] Sinha S and Pan R K 2005 *Preprint physics/0504198*