Self-organized coordination in collective response of non-interacting agents: Emergence of bimodality in box-office success

Anindya S. Chakrabarti\textsuperscript{1} and Sitabhra Sinha\textsuperscript{2}

\textsuperscript{1}Department of Economics, Boston University, 270 Bay State Road, Boston, MA-02134, USA.
\textsuperscript{2}The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113, India.

(Dated: September 16, 2013)

Many socio-economic phenomena are characterized by the appearance of a few “hit” products having a substantially higher popularity compared to their often equivalent competitors, reflected in a bimodal distribution of success (response). Using the example of box-office performance of movies, we show that the empirically observed bimodality can emerge via self-organization in a model where agents (theatres) independently decide whether to adapt a new movie. The response exhibits extreme variability even in the absence of learning or communication between agents and suggests that properly timing the release is a key determinant of box-office success.

PACS numbers: 05.65.+b,87.23.Ge,64.60.-i,89.65.-s

Complex systems often exhibit non-trivial patterns in the collective (macro) behavior arising from the individual (micro) actions of many agents [1]. Despite the high degree of variability in the characteristics of the individuals comprising a group, it is sometimes possible to observe robust empirical regularities in the system properties [2–4]. The existence of inequality in individual success, often measured by wealth or popularity, is one such universal feature [5]. While agents differ in terms of individual attributes, these can only partly explain the degree of this inequality [6]. The outcomes often have a heavy-tailed distribution with a much higher range of variability than that observed in the intrinsic qualities. Apart from the well-known Pareto law for income (or wealth) [7, 8], other examples include distributions of popularity for books [9], electoral candidates [10], online content [11] and scientific paradigms [12].

Another form of inequality may be observed in distribution of outcomes having a strongly bimodal character. Here events are clearly segregated into two distinct classes, e.g., corresponding to successes and failures respectively. While such distributions have been reported in many different contexts, e.g., gene expression [13], species abundance [14], wealth of nations [15], electoral outcomes [16], etc., one of the most robust demonstrations of bimodality is seen in the distribution of movie box-office success [17]. Here success is measured in terms of either the gross income $G_0$ at the opening weekend or the total gross $G_T$ calculated over the lifetime (i.e., the entire duration that a movie is shown) at theaters. Fig. 1 (a-b) shows that both of these distributions constructed from publicly available data for movies released in USA during the period 1997-2012 [18, 19] are described well by a mixture of two log-normal distributions. Although the movie industry has changed considerably during this time, the characteristic properties of the distributions appear to remain invariant over the successive intervals comprising the period. The log-normal character can be explained by the probability of movie success being a product of many independent chance factors [20], and is indeed observed in the distribution of the opening income per theater $g_0$ [Fig. 1 (c)]. However, the clear distinction of movies into two classes in terms of their box-office performance (as indicated by the occurrence of two modes in the $G_0$ and $G_T$ distributions) does not appear to be simply related to their intrinsic attributes [19, 21]. The fact that bimodality is manifested at the very beginning of a movie’s life also suggests that the extreme divergence of outcomes cannot be attributed to social learning occurring over time as a result of diffusion of information about movie quality [22] (e.g., by word-of-mouth [23]). Thus, although there have been theoretical attempts to explain emergence of bimodality through interaction between agents [24], we need to look for an alternative explanatory framework.

In this paper, we present a model for understanding the collective response of a system of agents to successive external shocks, where the behavior of each agent is the result of a binary decision process independent of other agents. Even in the absence of explicit interaction among agents, the system can exhibit remarkable coordination, characterized by the appearance of a strong bimodality in its response. For the specific example of box-office success, as the bimodal nature of the gross income distributions appear to be connected to the fact that movies usually open in either many or very few theaters, we focus on explaining the appearance of a bimodal distribution for the number of theaters $N_0$ in which movies open [Fig. 1 (d)]. Inspired by recent models that reproduce the observed invariant properties of financial markets by considering agents that interact only indirectly through their response to a common signal (price) [25], our model comprises agents (theaters) that do not explicitly interact with each other but whose actions achieve coherence by the regular arrival of a global stimulus, viz., new movies being introduced in the market. By contrast, decoherence is induced by the uncertainty under which a decision is made on releasing a new movie. We show that...
these competing effects can result in the appearance of bimodality in the distributions of $N_0$, and consequently, $G_0$ and $G_T$, where the success of a particular movie cannot be simply connected to its perceived quality prior to release nor to its actual performance on opening. Under a suitable approximation, we have analytically solved the model and obtained closed form expressions for peaks of the resulting multimodal distribution that match our numerical results. An important implication of our study is that the box-office performance of a movie is crucially dependent on whether it is released close in time to a highly successful one, which supports the popular wisdom that correctly timing the opening of a movie determines its fate at box-office.

We consider a system comprising $N$ agents (theaters) subjected to external stimuli (entry of new movies into the market), that have to choose a response, i.e., whether or not to adapt a new movie, displacing the one being shown. At any time instant $t$, this decision depends on a comparison between the perceived performance of the new movie and the actual performance of the movie being shown at the theater [Fig. 2 (a)]. For simplicity, we assume that a single new movie is up for release at each time instant $t$, thus allowing each movie to be identified by the corresponding value of $t$. The state of a theater at any time is indicated by the identity of the movie it screens at that time [Fig. 2 (b)]. The performance of a movie $t'$ at time $t$ can be measured by its income per theater, $g'$, which is related to its opening value $g_0$ by a scaling relation $g' = g_0(t - t')^{\beta_s}$. This relation is partly inspired by the empirical observation [Fig. 1 (d), inset] that the weekly income per theater for a movie decays as a power law function of the number of weeks after its release.

As agents are exposed to similar information about a movie that is up for release, they can have a common perception about its performance, measured as its predicted income. Although for most results reported here, $\beta_s = 0$ for simplicity, we have explicitly verified that qualitatively similar results are observed for other values of $\beta_s$ including $\beta_s = -1$.

At any time $t$, an agent $i$ switches to the new movie if it decides that this move will result in a higher income.

FIG. 1: (color online). Empirical demonstration of bimodality in movie popularity measured in terms (a) opening income $G_0$ and (b) total lifetime income $G_T$ of movies in theaters over successive intervals from 1997-2012 (indicated by different symbols). The data are fit by superposition of two log-normal distributions (broken curve). The cumulative distribution of the opening income per theater $g_0 = G_0/N_0$ over the same period is shown in (c). A fit with log-normal distribution is also indicated (broken curve). (d) The bimodal character of (a) and (b) can be connected to the bimodality observed in the distribution of the number of opening theaters $N_0$ (i.e., the total number of theaters in which a new movie is released). The inset shows the distribution of exponents $\beta$ characterizing the power-law decay of the weekly income per theater ($g_t \sim g_0 t^\beta$) for all movies.

FIG. 2: (color online). (a) Schematic diagram of the stochastic decision process of agents (theaters $i$, $j$ and $k$) who can either continue with “old” (movie being shown) or switch to “new” (movie up for release) at any time instant $t$. The probability that an agent $i$ will adopt the new movie, $p_{i,t}$, depends on a comparison of the perceived performance of that movie, $\theta_i$, to the actual performance of the movie being shown (which is related to its opening income $g_0$). (b) Time-evolution of a system comprising $N = 50$ agents (theaters), the state of each agent at any time being the movie (colored according to the time of release) that it is showing. At any time $t$, an agent $i$ switches to the new movie if it decides that this move will result in a higher income.
sufficient net gain $z_t(i) = \theta_t - g_t(i)$. It is implemented here by representing the probability of adapting the new movie as a hyperbolic response function [26]:

$$p_{i,t} = \frac{z_t(i)}{C + z_t(i)} \text{, for } z_t(i) \geq 0, \text{ else } p_{i,t} = 0, \quad (1)$$

where the parameter $C$ is an adaption cost incurred for switching to a new movie. We have verified that introducing more complicated functional forms for the adaption rule, e.g., ones having sigmoidal character, do not qualitatively change the results reported here. Eq. 1 allows us to calculate the number of opening theaters $N_0$ for every new movie [Fig. 2 (b)]. To obtain the opening income $G_0$ of the movie over all theaters that release it, $N_0$ is multiplied with the opening income per theater that is chosen from the log-normal distribution of $g_0$ referred to earlier [Fig. 1 (c)]. The subsequent decay of income per theater follows the empirical scaling relation with exponent $\beta$ [17]. The total lifetime income of a movie $G_T$ is obtained by aggregating this income for all theaters it is shown in, over the entire lifespan (i.e., from the time it is released until it is displaced from all theaters). While $\beta = -1$ for the results reported here, we have verified that our observations do not vary considerably when $\beta$ is distributed over a range. For most simulations, we have chosen $N = 3000$, which accords with the maximum number of theaters in the empirical data [19]. However, to verify that our results are not sensitively system-size dependent, we have checked that qualitatively similar behavior is observed for $N$ up to $10^6$.

As seen from Fig. 3, the system of $N$ independent agents self-organize in the limit of low $C$ to generate a bimodal distribution in their collective response. A new movie is either adapted by a majority [corresponding to the upper mode of the $N_0$ distribution shown in Fig. 3 (a)] or a small fraction [lower mode] of the total number of theaters. This translates into bimodal distributions in the opening income $G_0$ and total lifetime income $G_T$ [Fig. 3 (b-c)], which qualitatively resemble the corresponding empirically obtained distributions (Fig. 1). Thus, our results suggest that the nature of box-office income distributions for movies can be understood as an outcome of the bimodal character of the distribution for the number of theaters that release a movie coupled with the unimodal log-normal distribution for the income per theater. As the adaption cost $C$ is increased, the two modes approach each other until, at a large enough value of $C$, a transition to unimodal distribution for the quantities is observed. With increasing $C$, theaters are less likely to switch to a new movie, so that the time-interval between two consecutive movie releases at a theater becomes extremely long. This weakens temporal correlations between the performance of movies being shown and that expected from new movies up for release. Thus, the decision to release each new movie eventually becomes an independent stochastic event described by an unimodal distribution. To emphasize that bimodality in total income $G_T$ is a consequence of the bimodal nature of the opening income, we show $G_T$ as a function of the lifetime $T$ in Fig. 3 (d). We observe a bifurcation in $G_T$ at higher values of $T$ indicating that movies having the same lifetime can have very different total income, a feature previously seen in empirical data [27].

To understand the appearance of multiple peaks in the distribution of collective response in the limit of low adaption cost, we observe that the system dynamics is characterized by two competing effects: (a) the stochastic decision process of the individual theaters tend to increasingly decorrelate their states, while (b) the occasional appearance of movies having high $\theta$, that are perceived by the agents to be potential box-office successes, induces high level of coordination in response as a majority of agents switch to a common state. This phenomenon of gradual divergence in agent states interspersed by sporadic “reset” events that largely synchronize the system allows us to use the following simplification of the model for an analytical explanation. As $C \rightarrow 0$, we can approximate Eq. (1) by $p_{i,t} = p$ for $z_t(i) \geq 0$, else $p_{i,t} = 0$, which becomes accurate in the limit $p \rightarrow 1$. Thus, when a reset event occurs, the decision of each agent is a Bernoulli trial with probability $p$, so that the number of theaters that adapt the new movie follows a binomial distribution with mean $Np$ and variance $Np(1-p)$. In the limit $p \rightarrow 1$.
the number of opening theaters $N$ we obtain another peak at $N$ calculated as $\mu, \sigma$ parameters reset event to adapt the new movie with probability $\theta$. (a) The variations of opening income $G_0$ and total lifetime income $G_T$ of a movie as functions of the perceived performance $\theta$ and the actual performance (i.e., income per theater) $g_0$ shows that neither $\theta$ nor $g_0$ completely determine $G_0$ or $G_T (p = 0.9995)$. $
$ the variance becomes negligibly small and the distribution can be effectively replaced by its mean. This will correspond to a peak at $N_0^a = N p$, i.e., the higher mode. A movie that immediately follows a reset event can result in different responses from the agents depending on the value of $\theta$ associated with it. If this is larger than $g_0$ of all theaters, it is yet another reset event, the response to which is the same as above. However, if $\theta$ has a lower value that is nevertheless large enough to cause those theaters ($\approx N(1-p)$) which had not switched in the previous reset event to adapt the new movie with probability $p$, we obtain another peak at $N_0^i = N_p (1-p)$. This corresponds to the lower mode of the distribution. As seen from Fig. 4 (a), the two peaks of $N_0$ distribution are accurately reproduced by $N_0^a$ and $N_0^i$. In principle, the above argument can be extended to show that a series of peaks at successively smaller values of $N_0$ can exist at $N_p (1-p)^2$, $N_p (1-p)^3$, etc., but these will not be observed for the system size we consider here. The bimodal log-normal distribution of opening income $G_0$ results from a convolution of the multi-peaked distribution for $N_0$ with the log-normal distribution for $g_0$ (having parameters $\mu, \sigma$). The two modes of this distribution are calculated as $G_0^{\mu, \sigma} = \exp(\mu + \log N_0^{\mu, \sigma})$, which matches remarkably well with the numerical simulations of the model [Fig. 4 (b)].

While the individual behavior of agents are obviously dependent on the intrinsic properties (such as $\theta$) associated with specific stimuli, the collective behavior of the system cannot be reduced to a simple threshold-like response to external signals. Fig. 4 (c) shows that the opening income of different movies, which are segregated into two distinct clusters, are not simply determined by their perceived performance $\theta$, as one can find movies belonging to either cluster for any value of this quantity. Given that $\theta$ is only a prediction of the opening performance of a movie by the agents, and it need not coincide with reality, one may argue that the actual performance, i.e., the opening income per theater $g_0$, will be the key factor determining the aggregate income of the movie. However, Fig. 4 (d-e) show that neither the opening income nor the total lifetime income (both of which show clear separation into two clusters) can be explained as a simple function of the actual opening performance of the movie at a theater.

Our results explain box-office success as an outcome of competition between movies, where a new movie seeks to open at as many theaters as possible by displacing the older ones. Using an ecological analogy, a movie with high perceived performance invades and occupies a large number of niches until it is displaced later by a strong competitor. Thus, highly successful movies rarely coexist. This also implies that the response to a movie can be very different depending on whether or not it is released close to a reset event, i.e., the appearance of a highly successful movie (“blockbuster”). Therefore, our model provides explicit theoretical support to popular wisdom that timing the release of a movie correctly is a key determinant of its success at the box-office [28]. The critical importance of the launch time holds not only for movies, but also for many other short life-cycle products such as music, videogames, etc., whose opening revenues very often decide their eventual sales [29]. In fact, empirical data on movies show that for the dominant majority, the highest gross earning over all theaters they are shown in occurs on the opening weekend, followed by an exponential decay in income [17, 30].

To conclude, we have shown that extreme variability in response, characterized by a bimodal distribution, can arise in a system even in the absence of explicit interactions between its components. The observed inequality of outcomes cannot be explained solely on the basis of intrinsic variations in the signals driving the system. For a quantitative validation of the model we have used the explicit example of movie box-office performance whose bimodal distribution has been established empirically. Our analysis reveals that stochastic decisions on the basis of comparing effects of the preceding choice and the estimated impact of the upcoming movie gives rise to a surprising degree of coordination. The presence of bimodality in the absence of explicit interactions in several social

---

**FIG. 4:** (color online). Explaining the emergence of bimodal distribution in the limit of small adaption cost ($C \rightarrow 0$). The appearance of bimodality with parametric variation of the probability of adaption $p$ is shown for the distributions of (a) the number of opening theaters $N_0$ and (b) opening income $G_0$. As $p \rightarrow 1$, the approximation to the $C \rightarrow 0$ limit becomes more accurate. The results are obtained by averaging over many realizations with $N = 3000$ agents. The pair of thick lines in each figure indicate the theoretically predicted modes of the distributions (see text). (c-e) The variations of opening income $G_0$ and total lifetime income $G_T$ of a movie as functions of the perceived performance $\theta$ and the actual performance (i.e., income per theater) $g_0$ shows that neither $\theta$ nor $g_0$ completely determine $G_0$ or $G_T (p = 0.9995)$.
and biological systems suggests other possible applications of the theoretical approach presented here. Apart from bimodality, our model shows that more general multimodal distributions are possible in principle and empirical verification of this in natural and social systems will be an exciting development.

We thank Alex Hansen, Gautam I. Menon, Shakti N. Menon and Rajeev Singh for helpful discussions. This work was supported in part by IMSc Econophysics Project.

[30] In extremely few cases does a movie become more successful over time with its income exhibiting an increasing trend, eventually reaching a peak before again declining exponentially. To explain such rare “sleeper hits” [e.g., the movie My Big Fat Greek Wedding (2002) that achieved its highest gross around 20 weeks after its release], one may need to consider explicit interactions between agents.
Data description. We have studied income distribution (cross section) and time-path of incomes (time series) for movies released in U. S. A. within the span of last 16 years (1997-2012) for which reliable and exhaustive data is publicly available (http://www.the-movietimes.com). All income data are inflation-adjusted with respect to the price level of 2010 and if a movie is released in year \( x \), then both its opening and lifetime incomes are counted in year \( x \) even though it may run for more than one year. By income, we mean domestic income only (world-wide collection of income is ignored).

We have cross-sectional data on the gross opening income \( G_0 \) and lifetime income \( G_T \) from January, 1997 to July, 2012 which are divided into four windows viz., 97-00, 01-04, 05-08, 09-12 and the data is compressed across all years in each window to arrive at the bimodal cross-sectional income distributions in all windows and in aggregate (Fig. 1, a-b). The total number of movies released on which we have data is 4583 (=673 + 1240 + 1444 + 1226 in the four windows resp.) for \( G_0 \) and 5070 (= 1160 + 1240 + 1444 + 1226) for \( G_T \) (see table S1 below for year-by-year description). However, for the purpose of studying evolution of income of an individual movie in its life-time, we have data from July,1998 to July, 2012 and again the data-set is divided into four windows, the only difference being that now the first window has three years now (1998, 1999 and 2000) instead of four years. The total number of movies on which we have weekly data is 4568 (= 658 + 1334 + 1392 + 1184). Evidently the data set on the evolution of movie incomes is smaller than that for which we have only opening and final incomes. The gross opening incomes \( G_0 \) has been normalized by the number of movie theatres \( N_0 \) in which the movie was released to arrive at the distributions of per-theatre opening income \( g_0 \) which are mono-modal in all the four windows and in aggregate as well (Fig. 1, c).

All three aggregate variables \( N_0 \), \( G_0 \) and \( G_T \) are fitted with bimodal lognormal distributions (mixture of two lognormal distributions with weights \( \alpha \) and \( 1-\alpha \)) with parameters \( \alpha \), \( \mu_1 \), \( \mu_2 \), \( \sigma_1 \) and \( \sigma_2 \). The estimated values (maximum likelihood estimation) of the parameters of the \( N_0 \), \( G_0 \), \( G_T \) and \( g_0 \) are given in table S2. Next, we fit the time-path of incomes of movies over time assuming the following dynamics:

\[
\log(g_{t'}) = \text{constant} + \beta \log(t'-t+1) + \epsilon_{t'}
\]

where \( g_{t'} \) is the income per theatre of a movie released at time \( t \). We carried out the regression procedure over all movies that ran for at least 5 weeks and adjusted for missing reports of weekly income by excluding those weeks for which the income data is missing. The distribution of \( \beta \) clearly has a mode at -1 (Fig. 1, inset in (d)). This also establishes that the number of sleepers (movies that start with very low income but gain popularity in time) is negligibly small. Lastly, we have performed Hartigans’ dip test [? ] on our data and found that unimodality is rejected at the usual levels of significance.

The distribution of the maximum number of theaters \( N_{max} \) showing a particular movie during its entire lifetime of all movies in the period 1997-2012 with different symbols indicating different windows. Similar to the distribution of the number of opening theaters \( N_0 \) in Fig. 1-d, this also has a bimodal character with both modes being explained well by lognormal distributions (broken curve). (b) The probability distribution of the lifetime of all movies. The inset shows the fit of the cumulative distribution with a Weibull distribution. (c) The probability distribution of budget is unimodal and has positive correlation with lifetime income \( G_T \) (inset). Hence the source of bimodality in the opening/lifetime income of movies is not in budget. (d) Similar to the per theatre opening income in Fig. 1-c, the per theatre income of movies \( G_{w}/N_w \), for all weeks \( w \) over the entire lifetime is seen to be unimodal and is described well by a lognormal distribution.
SUPPLEMENTARY MATERIAL

<table>
<thead>
<tr>
<th>Year</th>
<th>N_0</th>
<th>G_0</th>
<th>G_T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>50</td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>50</td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>244</td>
<td>274</td>
<td>274</td>
</tr>
<tr>
<td>2000</td>
<td>294</td>
<td>299</td>
<td>299</td>
</tr>
<tr>
<td>2001</td>
<td>312</td>
<td>286</td>
<td>286</td>
</tr>
<tr>
<td>2002</td>
<td>370</td>
<td>321</td>
<td>321</td>
</tr>
<tr>
<td>2003</td>
<td>313</td>
<td>307</td>
<td>307</td>
</tr>
<tr>
<td>2004</td>
<td>339</td>
<td>326</td>
<td>326</td>
</tr>
<tr>
<td>2005</td>
<td>330</td>
<td>338</td>
<td>338</td>
</tr>
<tr>
<td>2006</td>
<td>344</td>
<td>354</td>
<td>354</td>
</tr>
<tr>
<td>2007</td>
<td>378</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>2008</td>
<td>340</td>
<td>368</td>
<td>368</td>
</tr>
<tr>
<td>2009</td>
<td>317</td>
<td>339</td>
<td>339</td>
</tr>
<tr>
<td>2010</td>
<td>354</td>
<td>358</td>
<td>358</td>
</tr>
<tr>
<td>2011</td>
<td>373</td>
<td>364</td>
<td>364</td>
</tr>
<tr>
<td>2012</td>
<td>140</td>
<td>165</td>
<td>165</td>
</tr>
<tr>
<td>Total</td>
<td>4568</td>
<td>5855</td>
<td>5855</td>
</tr>
</tbody>
</table>

TABLE S1: Number of movies analyzed for each year.

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β1</th>
<th>β2</th>
<th>σT</th>
<th>σ₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>N_0</td>
<td>0.61</td>
<td>2.91</td>
<td>7.84</td>
<td>1.72</td>
<td>0.29</td>
</tr>
<tr>
<td>N_max</td>
<td>0.64</td>
<td>4.01</td>
<td>7.83</td>
<td>1.71</td>
<td>0.27</td>
</tr>
<tr>
<td>G_0</td>
<td>0.57</td>
<td>11.36</td>
<td>16.49</td>
<td>1.24</td>
<td>0.94</td>
</tr>
<tr>
<td>G_T</td>
<td>0.54</td>
<td>13.16</td>
<td>17.55</td>
<td>1.80</td>
<td>1.05</td>
</tr>
<tr>
<td>g₀</td>
<td>5.72</td>
<td>8.72</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE S2: Values of the parameters estimated by maximum likelihood procedure.

The total number of theaters (agents) was assumed to be 3000 which is close to the maximum number of theaters in which a movie released in the time period considered (see distribution of N_max in Fig. ??). Each simulation was run for 5000 time periods and the first 500 periods were ignored to remove the effects of initial conditions. Then the results were averaged over multiple realizations to arrive at the steady state distributions. As is mentioned in the main text, the probability of release of each movie depends crucially on the factor z_t(i) = θ_t - g(i). Since we have assumed θ_t to be lognormally distributed with the estimated μ and σ of g₀ (see Table S2), we normalize z_t(i) by the mean of the distribution viz. exp(μ + σ²/2). In the models presented in the text, we have assumed β_t to be completely uncorrelated with g₀ and that allowed us to consider the simplest case, β_s = 0. The idea underlying the assumptions is that the agents make completely wrong predictions about movies’ performances and when the movie releases they come to know the actual performance. However, when they decide whether to adapt a movie or not, they compare the expected opening income of the upcoming movie to the realized opening income of the movie currently being shown. Given the monotonic decay of per theater weekly income, the opening income itself determines the future per theater weekly incomes. Thus, the agents are willing to adapt the new movie if and only if they perceive that the lifetime income of the new movie would be higher than the preceding movie. We have explored other combinations as well with qualitatively similar results as have been mentioned below.

SIMULATIONS

Unfortunately, such detailed database is available only for Hollywood movies released in America. Hence, we do not know if similar features are seen in other countries as well. However, these features seem to be invariant over a considerably long time horizon during which the American movie industry have undergone a lot of changes from the demand side (changes in taste and preference of movie viewers and technological improvements like introduction of Youtube, Netflix etc.) as well as the supply side (technological breakthroughs movie making, market-

ing and selling). Also it should be mentioned that the database we considered contains data on top 60 movies every week irrespective of genre or MPAA rating and includes movies with a staggering range of income and budget which reflects the fact that bimodality is not a local, temporary feature. And there is no evidence of bimodality in the movie budgets either that can explain the observations (Fig. ??). Lastly, though it is possible that the advent of social networking sites like Facebook, twitter etc. have an effect on polarization of opinion, that cannot be an explanation in the current context since the features were there even before Facebook or twitter were launched.

ROBUSTNESS

To check robustness, we simulated the same model by changing assumptions one by one. Assuming the agents
compare one month’s (four weeks/time points) expected earnings of the currently running and the new movies, the qualitative results did not differ much. In the baseline models, we assumed that the perceived and the actual income of the movies are completely uncorrelated and after releasing the new movie, agents come to know the true value of the opening income $g_0$. If the perceived and the actual values are equal (that is the agents make perfect forecasts about the movie’s performance), perhaps not so surprisingly, bimodality again appears naturally for some values of the cost parameter with both types of transition rules. Similar results hold if we add small noise to the actual $g_0$ (which was assumed to be a constant across theaters at every point of time) across the theatres e.g. if the realized opening income in each theater has a distribution with mean equal to $g_0$ or if we distribute the decay coefficient of income $\beta$ with a mean at -1 as is found in the data. Lastly, the model produces qualitatively similar results with another type of probabilistic choice function viz,

\[
p = 0.5 + 0.5 \left( \frac{z_t(i)}{\sqrt{C^2 + z_t(i)^2}} \right).
\]

(2)

Note that this function is smooth in $z$ whereas the other two mentioned in the main text were not so. This shows the robustness of the model with respect to different types of choice functions.