Taming the Unpredictability of Cultural Markets with Social Influence

Andrés Abeliuk
Data61 & MIT Media Lab
aabeliuk@mit.edu

Gerardo Berbeglia
Melbourne Business School
g.berbeglia@mbs.edu

Pascal Van Hentenryck
University of Michigan
pvanhent@umich.edu

Tad Hogg
Institute for Molecular Manufacturing
tadhogg@yahoo.com

Kristina Lerman
USC Information Sciences Institute
lerman@isi.edu

ABSTRACT

Unpredictability is often portrayed as an undesirable outcome of social influence in cultural markets. Unpredictability stems from the "rich get richer" effect, whereby small fluctuations in the market share or popularity of products are amplified over time by social influence. In this paper, we report results of an experimental study that shows that unpredictability is not an inherent property of social influence. We investigate strategies for creating markets in which the popularity of products is better—and more predictably—aligned with their underlying quality. For our study, we created a cultural market of science stories and conducted randomized experiments on different policies for presenting the stories to study participants. Specifically, we varied how the stories were ranked, and whether or not participants were shown the ratings these stories received from others. We present a policy that leverages social influence and product positioning to help distinguish the product's market share (popularity) from underlying quality. Highlighting products with the highest estimated quality reduces the "rich get richer" effect highlighting popular products. We show that this policy allows us to more robustly and predictably identify high quality products and promote blockbusters. The policy can be used to create more efficient online cultural markets with a better allocation of resources to products.

1. INTRODUCTION

Every day people make a staggering number of choices about what to buy, what to read, where to eat, and what to watch. The interplay between individual choices and collective opinion is responsible for much of the observed complexity of cultural markets and social behaviors. Predicting collective outcomes, such as the commercial success of movies and books is extremely difficult, even for experts [6, 8, 9, 12]. In the film industry, as an example, a few rare successes dominate box-office receipts [9]. Similar trends are evident online, where a few hit videos, songs, or Wikipedia pages receive the lion's share of attention, while most of the other products are barely noticed, a behavior typical of long-tailed distributions [27].

While popularity (or product's market share) is generally considered a mark of quality [26], the "irrational herding" effect created by social influence can obscure and distort perceptions of the underlying value of products. The MusicLab experiment [19], and multiple follow-up studies [20, 17, 24], demonstrated that social influence leads to *unpredictable* markets with gross *inequalities*, where a few of the products become vastly more popular than the rest, although it is difficult to predict exactly which products will become popular. Essentially, the information about the preferences of others—conveyed via social signals, such as sales volumes, box-office revenues, ratings on product review sites, and consumer recommendations—influences consumers' choices of products. In markets where products are ranked according to collective preferences—for example, in best seller lists, top-40 charts, hot lists, etc.—the combination of social influence and ranking order creates a feedback loop that leads to a "rich get richer" phenomenon [16]. This effect amplifies initial random fluctuations in popularity and decouples it from the underlying quality of products, making market outcomes both unpredictable and unequal.

Together, these findings paint a bleak picture for cultural markets. Unpredictability and inequality of product market share suggest that some high-quality products may go unnoticed [19, 21] on the one hand, while lower-quality ones may be oversold [11, 17] on the other. In either case, the market share does not entirely reflect user preferences, which may lead to unsatisfied consumers and loss of demand [20]. An important open research question is whether it is possible to mitigate the "rich get richer" effects in cultural markets so as to better align products' popularity with their underlying quality.

Several lines of work come to bear on this question. First, Duncan Watts [26] suggested the use of "measure and react" strategies to address the difficulty in making correct predictions about social behavior: "Rather than predicting how people will behave and attempting to design ways to make customers behave in a particular way [...] we can instead measure directly how they respond to a whole range of possibilities and react accordingly" [26]. Lerman and Hogg [16]

WWW 2017, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052680

showed that it is possible to steer collective outcomes toward a desired goal through the presentation order of items. Specifically, ordering items by recency of recommendation decreased the inequality and increased predictability of collective outcomes of peer recommendation [16]. Social influence, while increasing inequality, can reduce the amount of effort recommenders put into searching for products, increasing collective efficiency of the market [13]. Along similar lines, Abeliuk et al. [1], Van Hentenryck et al. [25] analyzed theoretically different policies for displaying products using the generative model of the MusicLab [14]. The authors of the study proved that under some policies the market converges almost surely to a monopoly for the product of highest quality, making the market both predictable and asymptotically optimal. Furthermore, their computational experiments confirm that some simple policies quickly identify blockbusters and outperform ranking products by their popularity, a policy which is ubiquitous in many real life applications, as well as in the MusicLab study.

In this paper, we experimentally investigate strategies to mitigate the "rich get richer" effects that emerge in cultural markets so as to more robustly identify blockbusters by better aligning the popularity of products with their underlying quality. To that end, we created an online trial-offer market consisting of a web interface displaying science news articles that participants can read and rate. Our web-based experiments consisted of 1,621 participants. In trial-offer markets, consumer choice is decomposed into two stages: a sampling stage where participants decide which item to try, followed by a second stage where participants decide whether to purchase/recommend the sampled product. Such trial-offer markets, where participants can try products before buying them, are pervasive in online cultural markets (e.g., books on Amazon, songs on iTunes and phone apps with free trial in Google Play). We evaluate different policies for presenting items to people and measure their impact on the efficiency and unpredictability of markets. The presentation policies use different criteria for ranking the items, and they may in addition use social influence by showing how others have rated the items. In addition to the *popularity ranking*, which puts more highly rated items above others, we also study the *quality ranking*, with and without social influence. Quality ranking orders items in the decreasing order of their quality, which is a dynamic quantity that is estimated from the observed actions and ratings of participants up to that point.

We show that social influence can be used to make markets more predictable and efficient. We find that, regardless of the ranking policy, social influence creates a "rich get richer" phenomenon that increases the inequality of market outcomes by creating very popular blockbusters at the expense of other items. Despite this, using the quality ranking with social influence improves the predictability of outcomes, compared to the popularity ranking. It leads to statistically similar levels of unpredictability as ranking policies that do not use social influence. Put together, these findings suggest that quality ranking with social influence is able to consistently push higher quality items to become blockbusters. These results contrast with conclusions of Salganik et al. study: it is not social influence per se that makes markets unpredictable, but the way it is used that leads to unpredictability.

2. RELATED WORK

The impact of social influence on collective behavior and market outcomes has a long history of study.

The MusicLab study created an artificial music market to experimentally investigate the impact of social influence [19]. Participants in the MusicLab study were presented a list of unknown songs where they had to decide which song to listen and after listening to a song, the participant had the opportunity to download it. The participants were divided into two groups exposed to two different experimental conditions: the independent condition and the social influence condition. In the independent condition group, participants were provided with no additional information about the songs. In the social influence condition group, each participant was provided with the number of times the song was downloaded by previous participants.

The MusicLab experiments relied on an implicit but critical design choice: songs were displayed to participants in decreasing order of popularity, reinforcing the social signal with position bias. As a result, the implicit (through ranking) and explicit (through displayed number of downloads) signaling of the preferences of others created synergies in participant's choices of songs to download. Leman and Hogg [16, 13] conducted a follow-up set of experiments, inspired by the MusicLab study design, to explicitly measure the contribution of position and social influence to item popularity. They found that position bias [18, 7] plays an important role in such cultural markets: people allocate significantly more visual attention to items appearing near the top of a Web page or a list of items than those below them. As a result, item ranking can create as much inequality and unpredictability as was observed in the MusicLab study, even in the absence of social influence. In fact, they found that "social influence affects popularity about half as much as position and content do" [13].

In the field of sponsored search, Agarwal et al. [3] show that the position of ad placement on clicks has a positive effect on click-through rates, but not necessarily in conversion rates. In the context of hotel bookings, Ghose et al. [10] experimentally studied the effect of ranking on revenue for different search engine policies. They propose a ranking that achieves higher search engine revenue compared with other mechanisms such as rankings based on price or ratings.

2.1 The MusicLab Model

The generative model of the MusicLab study [14] is based on data collected during the actual experiments and is accurate enough to reproduce the conclusions of Salganik et al. [19] through simulation. The model can be defined as a trial-offer market, where each participant is presented with a ranking of songs and each position in the ranking is characterized by its *visibility* v_j, which is the inherent probability of sampling a song in position j. For the first stage, the model specifies the probability of sampling song i given ranking σ as

$$p_i(\sigma) = \frac{v_{\sigma_i}(a_i + d_i)}{\sum_{j=1}^{n} v_{\sigma_j}(a_j + d_j)},$$

where each song is characterized by two values: its *appeal* a_i which represents the inherent preference of listening to song i based only on its name and its band; its download count d_i, which was display to participants as a social signal. For the second stage, each song has a *quality* q_i, which represents

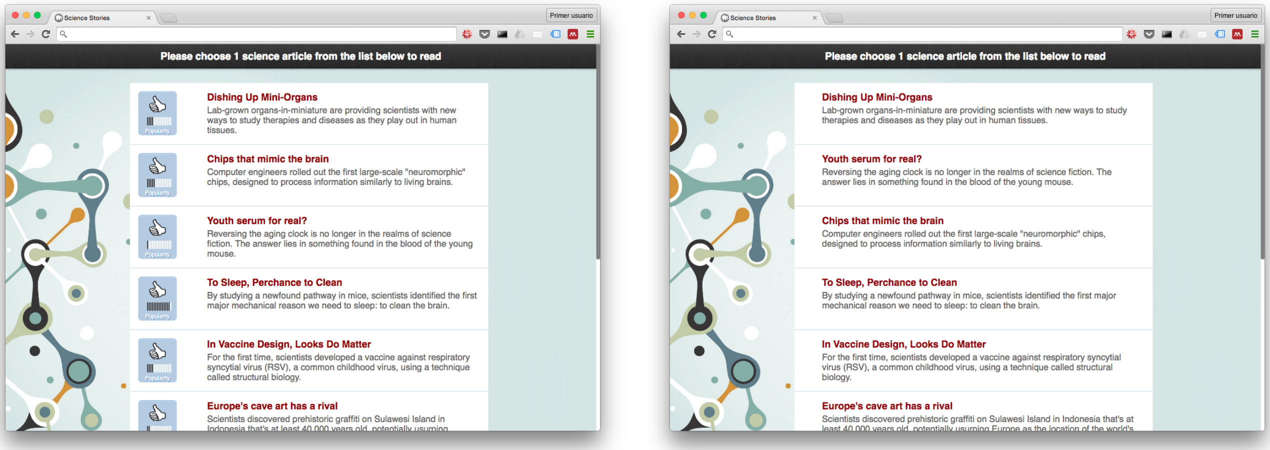

Figure 1: Screenshots of the web page shown to participants. The left picture is a screenshot of the story menu for the quality ranking when social signals are displayed. The right picture is a screenshot of the story menu for the random ranking when no social signals are displayed.

the conditional probability of downloading song i given that it was sampled.

This model was used in Abeliuk et al. [1, 2] to analyze a new policy for displaying the products, referred to as *performance ranking*. This ranking is a myopic policy that maximizes the efficiency of the market for each incoming participant, taking into account the inherent quality of products, position bias, and social influence. Computational and theoretical results showed that performance ranking significantly decreases the unpredictability of the market, and that the market reaches a unique monopoly for the highest quality product.

Performance ranking, however, presents challenges at the moment of implementation: one must have good approximations of the quality of items, appeal and visibility parameters of the model. A different policy, referred to as *quality ranking*, was proposed by Van Hentenryck et al. [25] to avoid estimating too many parameters of the model. Quality ranking only takes into account the inherent quality of products and yet, it produces results which are very similar to the performance ranking in terms of efficiency. Furthermore, product qualities can be recovered accurately and quickly, either before or during market execution.

Our experiments were designed specifically to test whether *an adaptive ranking of items in the presence of social signals can decrease the unpredictability of the market.* In particular, we designed the experiment to be close enough to the setting of trial-offer markets for quantitative comparisons with prior analytical results on the quality ranking [25].

3. METHODS

Our experiments showed participants a list of ten science stories displayed in a column and asked them to read one story and later recommend it if they found it interesting. To make the experiment appealing and to attract a large number of participants, the stories correspond to a subset of press releases, written for a wide audience, gathered from the top scientific breakthroughs list of Science Magazine from the years 2013 and 2014. Science news were selected given their longer lifespan compared to other types of news articles that only have a lifespan of a few days.

The participants were assigned (uniformly at random) into one of four different experimental conditions that vary depending on how the stories are ordered and whether social signals are displayed. If no social signals were present, then participants saw only story titles and short abstracts. When social signals were displayed, each participant was provided with additional information in the form of the number of recommendations that each story received from prior participants in that experiment. As a social signal, we used "Popularity bars", common in online music markets, such as iTunes and Spotify. The popularity bar shows the relative number of recommendations the story received. The most popular story in the list has the full bar, while the size of "popularity bars" of other stories is proportional to the number of recommendations the story has relative to the top story. Figure 1 illustrates the two different experimental conditions. After clicking on a story from the list, participants were able to view its full text and recommend it, if they chose. Figure 2 shows a screenshot of this step. After this step, users were asked to complete a short demographic survey, concluding their participation in the experiment; consequently, users were able to recommend at most one story. Table 2 (appendix) shows the demographic survey results. In addition, we used a "parallel worlds" design, which allowed us to rerun the experiment from the same initial conditions to compare the outcomes for the same experimental conditions.

Not all stories were equally interesting to people. When the same number of people view a story, a higher quality story receives more recommendations. Borrowing the idea of conditional quality from Krumme et al.'s model, we use the same measure as a proxy of quality, namely, the conversion rate. Formally, the conversion rate, $q_i = d_i/c_i$, is defined as the ratio of the number of recommendations the story received, d_i, to the number of times it was viewed, c_i.

The four experimental conditions (or policies) are listed below. For each policy, we created four "worlds" evolving completely independently.

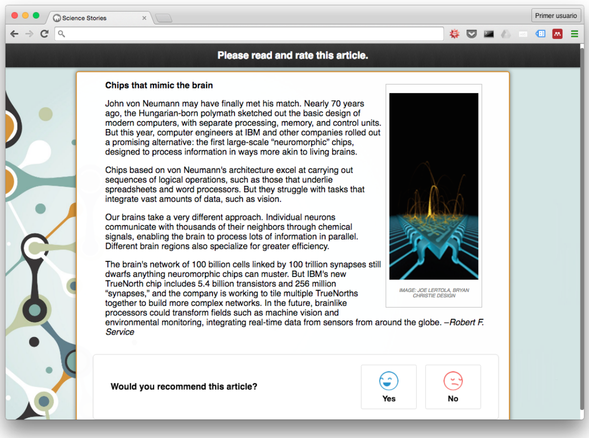

Figure 2: Screenshot of a science story page shown to participants. After clicking on a story from the story menu, users were prompt with the full story and had to either recommend or not the story in order to proceed. This step was the same for all policies.

Random ranking presented the stories in a new random order for each participant, and no social signals were displayed.

Popularity ranking ordered stories by their popularity within each world. Stories were sorted in decreasing order of the number of recommendations received up to that point in that world, and the social signals were displayed.

Quality ranking ordered stories in the decreasing order of their quality (conversion rate). The quality ranking was implemented with and without social signals.

- **Quality SI ranking** used the quality ranking with social influence: stories were sorted by quality and social signals were displayed;
- **Quality IN ranking** used the quality ranking under the independent condition: stories were sorted by quality and no social signals were displayed.

In order to more accurately approximate quality, we used the conversion rates from all worlds as the experiment progressed, combining them to determine the quality ordering for subsequent participants. The quality measure was updated with each action. Note that the quality ordering at any given moment of time is the same in all parallel worlds.

We created a total of 16 worlds, and among these worlds, four were of each policy. The study started with eight worlds, two for each policy, that were ran for two months in parallel. Two weeks after the beginning of the study, we launched eight additional worlds that operated in parallel for one and a half months. We refer to participants who successfully completed the task as "users" in our study. Table 1 shows a summary of the users in each of the worlds.

The University of Melbourne Human Ethics Advisory Group approved the experiment design. Our web-based experiments consisted of 1,621 participants which were recruited online via Google-sponsored ads. We created five versions of Google ads, each with slightly different wordings. An example of such an ad read as follows: "Top Science Stories: Help identify the scientific breakthroughs of the last years." Participants were able to perform the study after he or she agreed to the consent form specifying we were conducting a study of the role of social media in promoting science. Participation was unpaid and voluntary, making the experiment as close as possible to a generic social web platform.

Ranking Policy	Social Influence	World 1	2	3	4	Total
Popularity	Yes	104	108	101	94	407
Quality SI	Yes	103	104	102	97	406
Quality IN	No	103	107	104	93	407
Random	No	103	105	100	93	401
Total		413	424	407	377	1,621

Table 1: Summary of users for each world and ranking policy. Demographics about these participants are presented in the appendix.

4. RESULTS

4.1 Market unpredictability

To measure market unpredictability, we compare outcomes across different worlds using the same policy to present stories to users. Figure 4 depicts the unpredictability of outcomes using the measure proposed by Salganik et al. [19]: the unpredictability u_i of story i is defined as the average difference in the market share for that story for all pairs of worlds:

$$u_i = \sum_{w=1}^{W} \sum_{w'=w+1}^{W} |m_{i,w} - m_{i,w'}| / \binom{W}{2}.$$

Here, market share $m_{i,w}$ of story i in world w is $m_{i,j} = d_i^j / \sum_{k=1}^{n} d_i^j$, where d_i^w is the number of recommendations story i received in world w. The overall unpredictability is the average of this measure over all n stories, i.e., $U = \sum_{j=1}^{n} u_i / n$.

To estimate a 95% confidence interval of the unpredictability measure for each ranking policy, we calculate the unpredictability measure for all $\binom{4}{2} = 6$ possible pairs of worlds for that policy. The mean unpredictability value of all pairs is the same value as the overall unpredictability measure of the policy described above. We use the unpredictability measures based on pair of worlds to calculate a 95% confidence interval using the bootstrapping approach. We performed the Mann-Whitney test, which makes a pairwise comparison between two sets of unpredictability measures based on pair of worlds, one for each policy, to determine statistical significance of results.
Figure 4 highlights two interesting results:

1. In the presence of social influence, quality ranking leads to significantly (p-value$<$ 0.01) less unpredictability than the popularity ranking.
2. The quality ranking with social influence has similar levels of unpredictability compared to policies that are

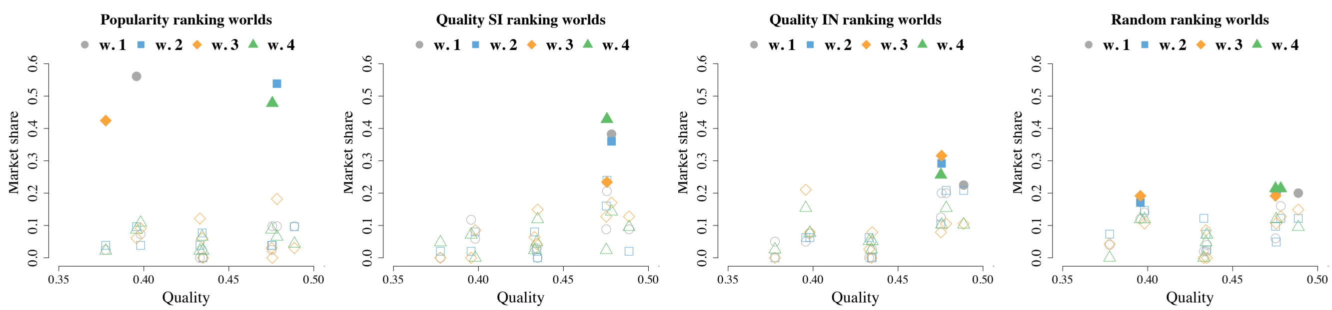

Figure 3: The distribution of recommendations market share for each world. Each dot is the recommendations market share of a story in a world vs the quality measure of that story. The most recommended stories per world are highlighted in dark. If ties exist, we highlighted all the stories that share the first position.

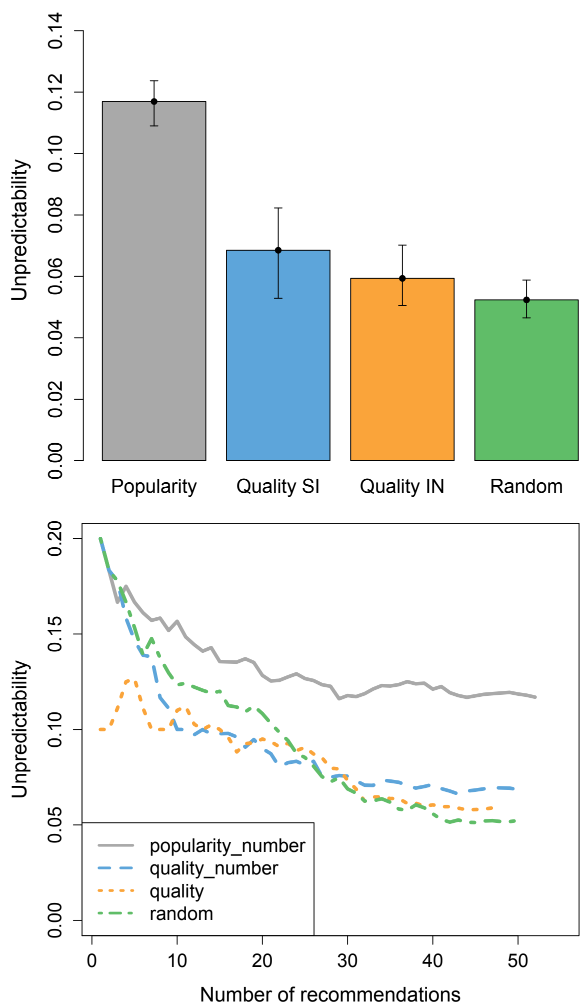

Figure 4: Unpredictability of outcomes for different ranking policies. (Top) The overall unpredictability for each policy. Error bars indicate 95% confidence intervals. (Bottom) The dynamics of unpredictability as the experiment progresses. The final values of each policy correspond to the values on the top graph.

not using social signals. At a 0.05 significance level, we cannot conclude that the quality ranking with social signals has a greater unpredictability than the quality ranking without social influence (p-value$=$ 0.20). The same holds when comparing the quality ranking with social signals against the random ranking (p-value$=$ 0.066). Given the small standardized mean difference between the unpredictability of these policies, our study was insufficiently powerful to detect them. If differences do exist, the sample size (number of worlds) needed to reach statistical significant difference, based on the size of the difference we see in our study, is calculated with a power analysis. In particular, 50 worlds per policy are needed to compare the quality ranking with social signals against the quality ranking without social influence, and 16 worlds per policy to compare the quality ranking with social signals against the random ranking.

The intuition behind why the popularity ranking has the highest level of unpredictability is that the stories were displayed in decreasing order of popularity, reinforcing the social signal with position bias and hence, leading to a "rich get richer" effect.

It has been argued that social influence makes markets unpredictable [19] and as a result, social influence is often presented in a negative light. Our experimental results show that unpredictability is not an inherent property of social influence. Whether a market is predictable or not depends on how social influence is used.

4.2 Emergence of blockbusters

Social influence and item ranking create a "rich get richer" phenomenon that can obscure the underlying value of products. As a result, even average quality products can become blockbusters, while some gems are overlooked [19]. Do blockbusters arise for the quality ranking?

Figure 3 reports the distributions of recommendations market share of every story in every world compared to the quality measure of that story across all worlds. The figure highlights three interesting results:

1. Unambiguous blockbusters emerge when social influence exists: there is a clear gap between the most popular story and the rest of the stories. The gap decreases as we move to the worlds with no social signals, and it is smallest for the random ranking policy. The most

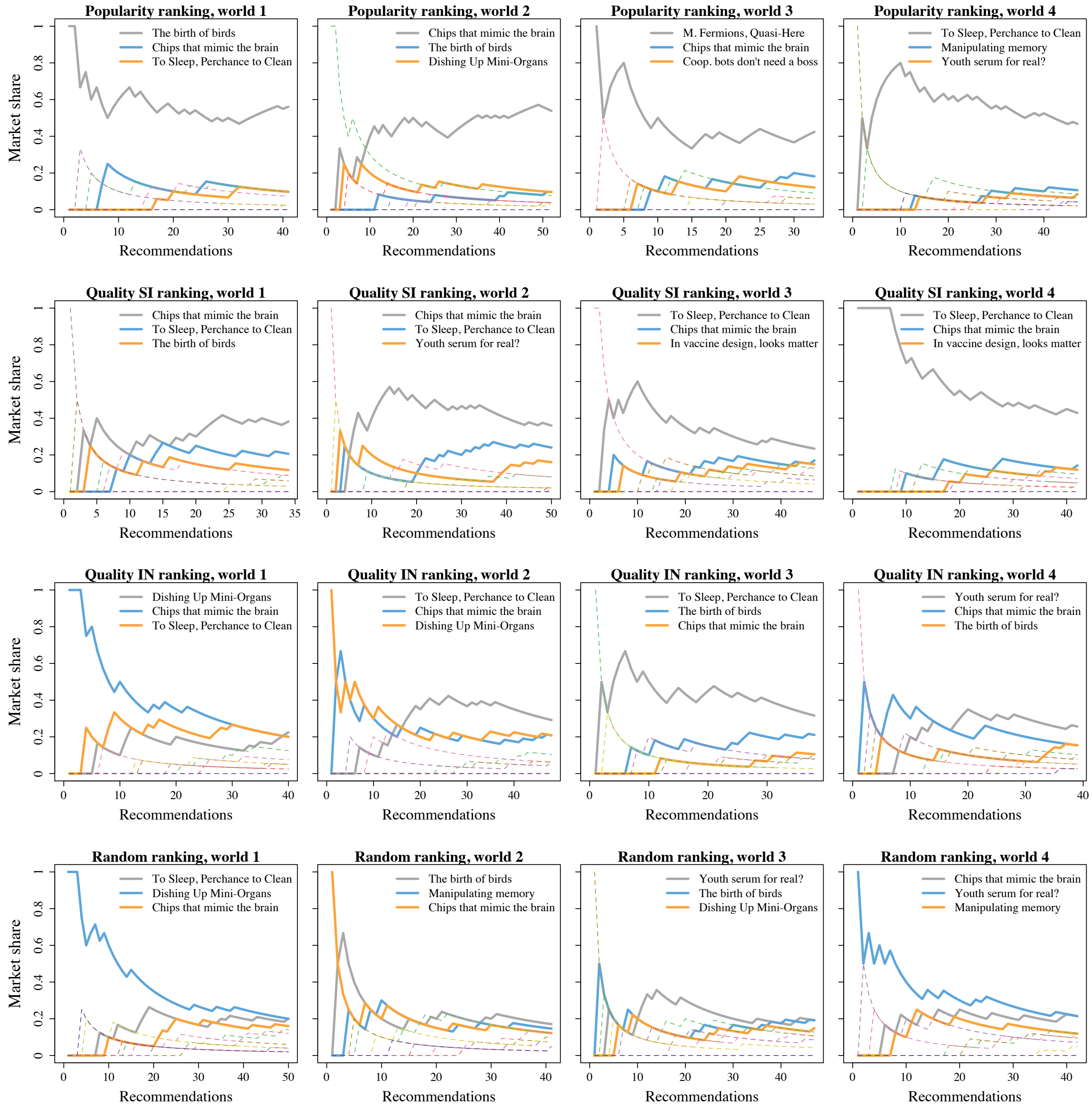

Figure 5: Market share dynamics of all stories in each world. Continuous bold lines represent the three most recommended stories of the world.

popular stories on average accrued 50%, 35%, 27%, 19% of the total recommendations for the popularity ranking, quality SI ranking, quality IN ranking and random ranking, respectively. Inequality is most pronounced when position bias and social influence align, that is when social signals exist and the top-ranked stories are ones with the largest social signals.

2. Most of the unpredictability that arises in worlds with social influence can be explained by the diversity of the most recommended stories in each world (highlighted in Figure 3). For the popularity ranking, each of the worlds gave rise to a different most recommended story. Whereas for the quality ranking, two stories—similar in terms of quality—were the most recommended in all four worlds.

3. The popularity ranking may lead to "low-quality" stories becoming blockbusters, which only occurred in worlds 1 and 3 for the popularity ranking.

These results shed light on the nature of social influence. First, the popularity ranking shows that social influence can turn average stories into blockbusters. However, this behavior is entirely eliminated by the quality ranking that leverages social influence to isolate blockbusters, consistently resulting in high quality stories becoming the most-recommended stories. The contrast between the distribution of popularity with and without social influence are

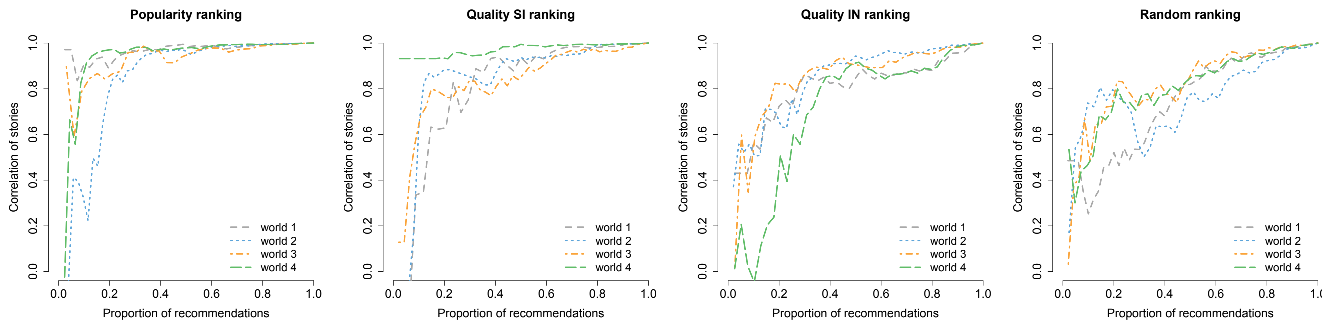

Figure 6: Pearson's correlation coefficient with the final distribution of popularity in each world as experiment progresses.

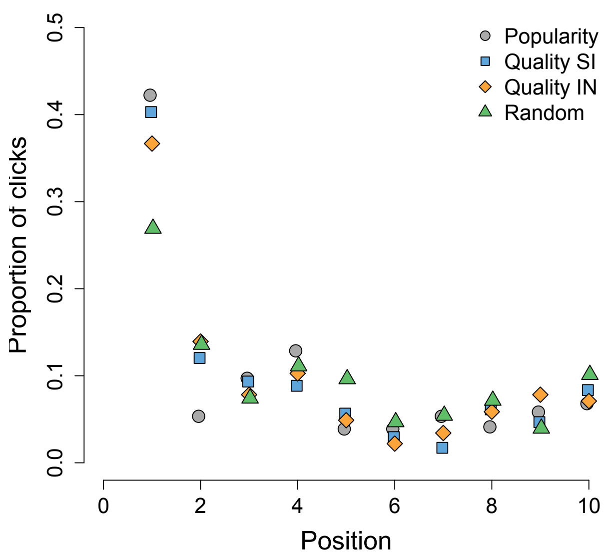

Figure 7: Position bias: proportion of users clicking on stories displayed in different positions. Each dot is the overall proportion of clicks that stories got in that position across all worlds using the same ranking policy.

also worth noting: popular stories under social influence seem to exhibit extreme imbalance with few stories dominating the market. Such distributions are commonly seen in "rich get richer" models, such as the preferential attachment model [5].

4.3 First mover advantage

Since popularity is amplified by the "rich get richer" effect, even small initial fluctuations can lead to large difference in outcomes: what is known as the "first-mover" advantage. If this is the case, the first users determine the trajectory of the system and hence, the final outcome should be highly dependent on their recommendations.

Figure 5 depicts the market share of all stories for each world as a function of the number of recommendations. For the popularity ranking, in worlds 1 and 3 the first story recommended became the most popular. In worlds 2 and 4, the second most recommended story by the first five users became the most popular story. For the quality ranking with social influence, a similar dependency of the first recommendations can be seen, although to a lesser extent. For the policies not using social signals, there seems to be a tendency for the stories that were recommended first to get second place (blue lines); however, the most popular stories (gray lines) emerge after the first recommendations.

Next, we look at how quickly the distribution of popularity approaches its final value in each world. Figure 6 shows Pearson's correlation coefficient between the final distribution of the number recommendations and its distribution as the experiment progresses. Naturally, such a metric should converge to one (meaning complete correlation), so we look at the rate of convergence. The policies with social influence appear to converge faster than those without—in fact, high correlations can be seen in most cases after the first 10% of the recommendations are made. Comparing the popularity ranking with quality under social influence, there seems to be a weak tendency for the popularity ranking to have higher correlations, suggesting that early users control the dynamics. Clearly, the quality ranking is not exempt from this behavior as is the case in world 4. In this particular world, the story that was at the top in the beginning received many clicks and remained in the top position for the first quarter of the experiment.

4.4 Position bias

Figure 7 shows the proportion of clicks stories received when shown to users at each position within the list. This measure quantifies the likelihood that any story is clicked based only on its position, i.e., the visibility of the position. This value varies substantially: stories shown to users at the top of the list receive about 30% of all the clicks under the random ordering condition and around 40% under the social influence condition. This arises due to a cognitive bias known as position bias [18]: people tend to pay more attention to items appearing in top list positions than those below them. Social influence amplifies position bias. The top-ranked story receives more attention under a ranking policy that uses social signals than under one that does not use social signals.

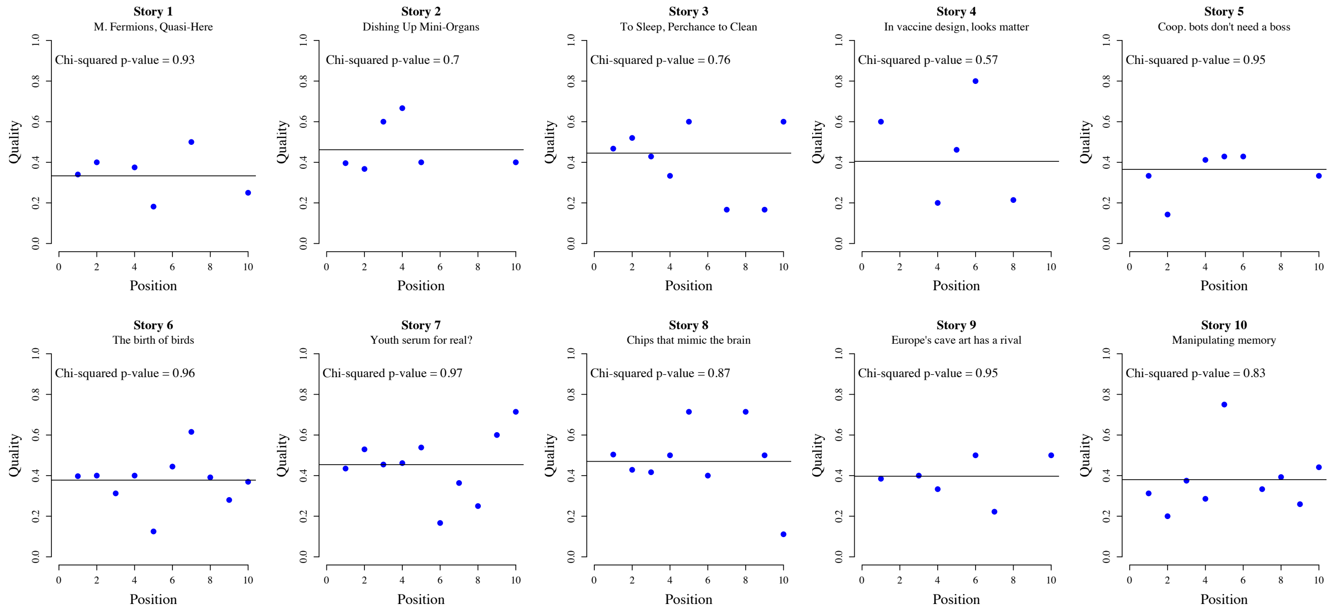

Figure 8: Average quality per story vs position across all worlds. The horizontal line is the average quality and the p-values correspond to the χ-squared test of homogeneity of proportions.

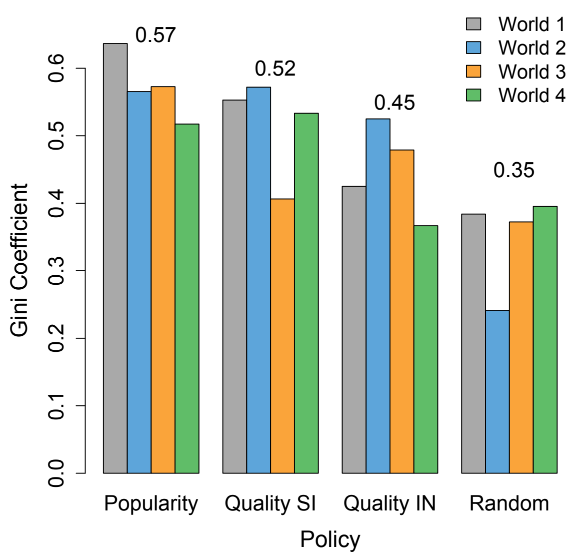

Figure 9: Inequality of popularity for different ranking policies, as measured by the Gini coefficient.

4.5 Market inequality

Figure 9 reports inequality of the market share of stories arising for different ranking policies. The inequality is measured in terms of the Gini coefficient [19]. Results show that policies using social signals create larger inequality than policies not using social signals. The inequality is most pronounced for the popularity ranking (*cf* Figure 3). This is likely due to the strong "rich get richer" effect caused by the alignment of position bias and social influence. As a result of this alignment, popular stories that appear in top list positions receive more attention than others stories, which leads them to receive more recommendations and become even more popular.

The quality-based ranking produces more equitable outcomes than the popularity ranking, though more than would be expected given variations in story quality. Adding social influence creates more inequality in the quality ranking, also likely due to the herding effect described above.

4.6 Perceived quality

The quality policy orders stories in decreasing order of their quality, where by quality we mean the conditional probability a user recommends a story after clicking on its url to view its full content. Since our experiments recorded when users clicked on a url to view a story, we can directly estimate quality from the data. The decision to recommend the story is made after the story is sampled from the list; hence, we hypothesize that position bias and social influence do not affect the quality of a story. This in line with past experiments in online trial-offer markets for music where after-consumption ratings of songs have low correlation with social influence [22]. Next, we explore if popularity or position biases affect perceived quality.

Figure 8 presents the estimated quality of each story, broken down by the position in which the story was displayed. The estimates are generated by aggregating observations over all worlds for the same story. We performed a χ-squared test of homogeneity for every story where the null hypothesis is that each position receives the same proportion of recommendations. For all stories, the p-values are higher than a 0.23 level of significance and hence, there is no reason to reject the null hypothesis. To test the independence of quality from social influence, we performed the χ-squared test of homogeneity for each story aggregating observations for all positions but broken down by the four policies. The p-values for each story (p-values: 0.87, 0.68, 0.92, 0.55, 0.67, 0.59, 0.64, 0.25, 0.84, 0.99) are high, leading to the conclusion that the ranking, whether it has social influence or not, does not affect the quality of stories. Our findings suggest that the effects of position bias and social influence have a negligible impact on the probability that a user recommends a story; biases have a greater effect on the decisions made during the sampling stage.

5. DISCUSSION AND CONCLUSION

Our results shed new light on the impact of social influence on cultural markets. While the unpredictability stemming from social influence is often presented as a strongly undesirable property of cultural markets, our experiments show that unpredictability is not inherent to social influence, but rather a consequence of the ranking policy used. Indeed, social influence can help make markets more predictable and efficient. In our online experiments, we showed that ranking products by quality under social influence had similar levels of unpredictability as ranking policies that did not use social influence. On the other hand, social influence created a "rich get richer" effect, evident in the large inequality of popularity of stories, which was not easy to counteract even when stories were ranked by quality. Taken together, these results suggest that quality ranking with social influence is able to consistently push higher quality stories to become blockbusters, creating a more efficient market. In contrast, the popularity-based ranking transforms average stories into blockbusters, creating market inefficiencies.

Our work has important limitations that suggest directions for future research. First, although the present study offers important insights about the role of ranking policies in the unpredictability of cultural markets, it does not provide strong statistical evidence that unpredictability can be completely mitigated. In future research it would be prudent to consider larger sample sizes and more parallel worlds per policy. Second, in practice, we do not see multiple worlds to estimate quality. Instead, we can only rely on the popularity of items in a single world, and any measure of quality in this version of the world could be biased by the past activity [13]. To circumvent this problem, quality could be estimated using methods that account for biases (for e.g. [23]). Lastly, our work studies one specific type of social influence; however, social signals may come in various forms and types. For example, advertisers are using methods that use information about consumers' social networks for recommendations in the form of personalized social signals that show the past activity of friends. Personalized social cues may have a greater impact than anonymized (aggregated) signals in the decisions of news reading [15] and advertising [4]. Disentangling the effects of position bias and the different types of social signals will lead to better ranking strategies to control and reduce the unpredictability of the market. Further experiments on this direction are left to future work.

We hope that these results will revive the debate about the consequences of social influence. Our results show that social influence can help detect correctly good quality products and that much of its induced unpredictability can be controlled. Our findings also suggest that the model in Krumme et al. [14] is a good policy making tool that helps analyze collective behavior in cultural markets under different policies [1, 2, 25] and contributes to continue building our understanding of the nature of social influence and to determine how best to use it.

Appendix

World	1	2	3	4
No. of participants	407	406	407	401
Age (%)				
Under 18 years	4.2	6.2	5.9	5.5
18-29 years	7.1	7.6	7.1	8.2
30-49 years	2.9	3.7	4.4	3
50-69 years	1.7	2.2	4.2	3.5
70 or older	2	3.4	2.7	1.7
N/A	82.1	76.9	75.7	78.1
Language (%)				
Arabic	0.7	1	0	0.5
English	8.7	13.1	16	14.0
Hindi	1.7	1.5	0.2	1.5
Mandarin	0.2	0.7	1	0.2
Portuguese	0.7	0.2	0.7	0.5
Spanish	2	1.5	1.2	1.5
Other	4.9	4.9	4.7	3.7
N/A	81.1	77.1	76.2	78.1
Education (%)				
Some high school	4.7	3.9	4.2	5.5
High school graduate	3.4	5.2	2.5	4.5
Some college	1.8	4.9	3.4	2.2
College graduate	5.4	4.4	5.2	3.8
Some postgraduate	1	1.2	2.9	2
Post graduate degree	3.4	2.5	5.2	4.7
N/A	80.3	77.8	76.7	77.3

Table 2: Summary of demographic questions from a survey we gave to participants at the end of their participation in the experiment. Users had the option to leave the survey's question blank. Not available (N/A) values apply for all participants who did not answer that question.

Acknowledgement

We would like to thank Manuel Cebrian for insightful discussions that greatly improved the design of the experiment. Work was partly funded by CSIRO's Data61, the ARO (W911NF-15-1-0142) and NSF (SMA-1360058).

References

[1] Abeliuk, A., Berbeglia, G., Cebrian, M., and Van Hentenryck, P. (2015). The Benefits of Social Influence in Optimized Cultural Markets. *PLOS ONE*, 10(4).

[2] Abeliuk, A., Berbeglia, G., Maldonado, F., and Van Hentenryck, P. (2016). Asymptotic Optimality of Myopic Optimization in Trial-Offer Markets with Social Influence. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16).*

[3] Agarwal, A., Hosanagar, K., and Smith, M. D. (2011). Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets. *Journal of Marketing Research*, 48(6):1057–1073.

[4] Bakshy, E., Eckles, D., Yan, R., and Rosenn, I. (2012). Social influence in social advertising: evidence from field experiments. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 146–161. ACM.

[5] Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–12.

[6] Bielby, W. T. and Bielby, D. D. (1994). All hits are flukes: Institutionalized decision making and the rhetoric of network prime-time program development. *American Journal of Sociology*, 99(5):1287–1313.

[7] Buscher, G., Cutrell, E., and Morris, M. R. (2009). What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 21–30. ACM.

[8] Caves, R. E. (2000). *Creative industries: Contracts between art and commerce*. Harvard University Press.

[9] De Vany, A. and Walls, W. D. (1999). Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office? *Journal of Cultural Economics*, 23(4):285–318.

[10] Ghose, A., Ipeirotis, P. G., and Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654.

[11] Hanson, W. A. and Putler, D. S. (1996). Hits and misses: Herd behavior and online product popularity. *Marketing letters*, 7(4):297–305.

[12] Hirsch, P. M. (1972). Processing fads and fashions: An organization-set analysis of cultural industry systems. *American journal of sociology*, pages 639–659.

[13] Hogg, T. and Lerman, K. (2015). Disentagling the effects of social signals. *Human Computation Journal*, 2(2):189–208.

[14] Krumme, C., Cebrian, M., Pickard, G., and Pentland, S. (2012). Quantifying social influence in an online cultural market. *PloS one*, 7(5):e33785.

[15] Kulkarni, C. and Chi, E. (2013). All the news that's fit to read: a study of social annotations for news reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2407–2416. ACM.

[16] Lerman, K. and Hogg, T. (2014). Leveraging Position Bias to Improve Peer Recommendation. *PLOS ONE*, 9(6):1–8.

[17] Muchnik, L., Aral, S., and Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science*, 341(6146):647–651.

[18] Payne Stanley, L. (1951). *The Art Of Asking Questions*. Princeton University Press.

[19] Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856.

[20] Salganik, M. J. and Watts, D. J. (2008). Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71(4):338–355.

[21] Salganik, M. J. and Watts, D. J. (2009). Web-Based Experiments for the Study of Collective Social Dynamics in Cultural Markets. *Topics in Cognitive Science*, 1(3):439–468.

[22] Sharma, A. and Cosley, D. (2013). Do social explanations work?: studying and modeling the effects of social explanations in recommender systems. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1133–1144. ACM.

[23] Stoddard, G. (2015). Popularity and quality in social news aggregators: A study of reddit and hacker news. In *Proceedings of the 24th International Conference on World Wide Web*, pages 815–818. ACM.

[24] Van de Rijt, A., Kang, S. M., Restivo, M., and Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 111(19):6934–6939.

[25] Van Hentenryck, P., Abeliuk, A., Berbeglia, F., Maldonado, F., and Berbeglia, G. (2016). Aligning Popularity and Quality in Online Cultural Markets. In *Proceedings of the tenth International AAAI Conference on Web and Social Media (ICWSM-16)*, pages 398–407.

[26] Watts, D. J. (2012). *Everything Is Obvious: How Common Sense Fails Us.* Random House LLC.

[27] Wilkinson, D. M. and M., D. (2008). Strong regularities in online peer production. In *Proceedings of the 9th ACM conference on Electronic commerce - EC '08*, pages 302–309. ACM.