Science AAAS

Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market

Matthew J. Salganik, *et al.*

Science **311**, 854 (2006);
DOI: 10.1126/science.1121066

Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market

Matthew J. Salganik,[1,2]* Peter Sheridan Dodds,[2]* Duncan J. Watts[1,2,3]*

Hit songs, books, and movies are many times more successful than average, suggesting that "the best" alternatives are qualitatively different from "the rest"; yet experts routinely fail to predict which products will succeed. We investigated this paradox experimentally, by creating an artificial "music market" in which 14,341 participants downloaded previously unknown songs either with or without knowledge of previous participants' choices. Increasing the strength of social influence increased both inequality and unpredictability of success. Success was also only partly determined by quality: The best songs rarely did poorly, and the worst rarely did well, but any other result was possible.

How can success in cultural markets be at once strikingly distinct from average performance (*1–4*), and yet so hard to anticipate for profit-motivated experts armed with extensive market research (*4–8*)? One explanation (*9*) for the observed inequality of outcomes is that the mapping from "quality" to success is convex (i.e., differences in quality correspond to larger differences in success), leading to what has been called the "superstar" effect (*9*), or "winner-take-all" markets (*10*). Because models of this type, however, assume that the mapping from quality to success is deterministic and that quality is known, they cannot account for the observed unpredictability of outcomes. An alternate explanation that accounts for both inequality and unpredictability asserts that individuals do not make decisions independently, but rather are influenced by the behavior of others (*11*, *12*). Stochastic models of collective decisions that incorporate social influence can exhibit extreme variation both within and across realizations (*4*, *13*, *14*), even for objects of identical quality (*3*, *15*). Unfortunately, empirical tests of these predictions require comparisons between multiple realizations of a stochastic process, whereas in reality, only one such "history" is ever observed.

We adopted an experimental approach to the study of social influence in cultural markets. We created an artificial "music market" (*16*) comprising 14,341 participants, recruited mostly from a teen-interest World Wide Web site (*17*), who were shown a list of previously unknown songs from unknown bands (*18*). In real time, arriving participants were randomly assigned to one of two experimental conditions—independent and social influence—distinguished only by the availability of information on the previous choices of others. In the independent condition, participants made decisions about which songs to listen to, given only the names of the bands and their songs. While listening to a song, they were asked to assign a rating from one star ("I hate it") to five stars ("I love it"), after which they were given the opportunity (but not required) to download the song. In the social influence condition, participants could also see how many times each song had been downloaded by previous participants. Thus, in addition to their own musical preferences, participants in the social influence condition received a relatively weak signal regarding the preferences of others, which they were free to use or ignore. Furthermore, participants in the social influence condition were randomly assigned to one of eight "worlds," each of which evolved independently of the others. Songs in each world accumulated downloads only from participants in that world, and subsequent participants could only see their own world's download counts.

Our experimental design has three advantages over both theoretical models and observational studies. (i) The popularity of a song in the independent condition (measured by market share or market rank) provides a natural measure of the song's quality, capturing both its innate characteristics and the existing preferences of the participant population. (ii) By comparing outcomes in the independent and social influence conditions, we can directly observe the effects of social influence both at the individual and collective level. (iii) We can explicitly create multiple, parallel histories, each of which can evolve independently. By studying a range of possible outcomes rather than just one, we can measure inherent unpredictability: the extent to which two worlds with identical songs, identical initial conditions, and indistinguishable populations generate different outcomes. In the presence of inherent unpredictability, no measure of quality can precisely predict success in any particular realization of the process.

We report the results of two experiments in which we study the outcomes for 48 songs by different bands (*18*). In both experiments, all songs started with zero downloads (i.e., all initial conditions were identical), but the presentation of the songs differed. In the social influence condition in experiment 1, the songs, along with the number of previous downloads, were presented to the participants arranged in a 16 × 3 rectangular grid, where the positions of the songs were randomly assigned for each participant (i.e., songs were not ordered by download counts). Participants in the independent condition had the same presentation of songs, but without any information about previous downloads. In experiment 2, participants in the social influence condition were shown the songs, with download counts, presented in one column in descending order of current popularity. Songs in the independent condition were also presented with the single column format, but without download counts and in an order that was randomly assigned for each participant. Thus, in each experiment, we can observe the effect of social influence on each song's success, and by comparing results across the two experiments, we can measure the effect of increasing the "strength" of the relevant information signal.

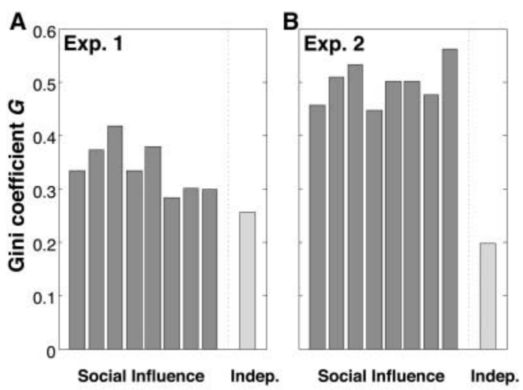

Fig. 1. Inequality of success for social influence (dark bars) and independent (light bars) worlds for (**A**) experiment 1 and (**B**) experiment 2. The success of a song is defined by m_i, its market share of downloads ($m_i = d_i / \sum_{k=1}^{S} d_k$, where d_i is song i's download count and S is the number of songs). Success inequality is defined by the Gini coefficient $G = \sum_{i=1}^{S} \sum_{j=1}^{S} |m_i - m_j| / 2S \sum_{k=1}^{S} m_k$, which represents the average difference in market share for two songs normalized to fall between 0 (complete equality) and 1 (maximum inequality). Differences between independent and social influence conditions are significant ($P < 0.01$) (*18*).

[1]Department of Sociology, 413 Fayerweather Hall, Columbia University, New York, NY, 10027, USA. [2]Institute for Social and Economic Research and Policy, Columbia University, 420 West 118th Street, 8th Floor, New York, NY, 10027, USA. [3]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM, 87501, USA.

*To whom correspondence should be addressed. E-mail: mjs2105@columbia.edu (M.J.S.); pd315@columbia.edu (P.S.D.); djw24@columbia.edu (D.J.W.)

Our results support the hypothesis that social influence, which here is restricted only to information regarding the choices of others, contributes both to inequality and unpredictability in cultural markets. Figure 1 displays the effects of social influence on market inequality, as measured by the Gini coefficient (*19*) (other measures yield similar results). In both experiments, we found that all eight social influence worlds (dark bars) exhibit greater inequality—meaning popular songs are more popular and unpopular songs are less popular—than the world in which individuals make decisions independently (light bars). Comparing Fig. 1, A and B, we also note that inequality increased when the salience of the social information signal was increased from experiment 1 to experiment 2. Thus our results suggest not only that social influence contributes to inequality of outcomes in cultural markets, but that as individuals are subject to stronger forms of social influence, the collective outcomes will become increasingly unequal.

Social influence also generates increased unpredictability of outcomes (Figs. 2 and 3). In each experiment, the average difference in market share (fraction of total downloads) for a song between distinct social influence worlds is higher than it is between different subpopulations of individuals making independent decisions (Fig. 2). Because these different outcomes occur even with indistinguishable groups of subjects evaluating the same set of songs, this type of unpredictability is inherent to the process and cannot be eliminated simply by knowing more about the songs or market participants. Figure 3 displays the market share (left column) and market rank (right column) of each song in each of the eight social influence worlds as a function of its "quality" (i.e., its market share and rank, respectively, in the independent condition). Although, on average, quality is positively related to success, songs of any given quality can experience a wide range of outcomes (Fig. 3). In general, the "best" songs never do very badly, and the "worst" songs never do extremely well, but almost any other result is possible. Unpredictability also varies with quality—measured in terms of market share, the "best" songs are the most unpredictable, whereas when measured in terms of rank, intermediate songs are the most unpredictable (this difference derives from the inequality in success noted above). Finally, a comparison of Fig. 3, A and C, suggests that the explanation of inequality as arising from a convex mapping between quality and success (*9*) is incomplete. At least some of the convexity derives not from similarity of preexisting preferences among market participants, but from the strength of social influence.

Our experiment is clearly unlike real cultural markets in a number of respects. For example, we expect that social influence in the real world—where marketing, product placement, critical acclaim, and media attention all play important roles—is far stronger than in our experiment. We also suspect that the effects of social influence were further diminished by the relatively small number of songs, and by our requirements (which aided control) that subjects could participate only once and could not share opinions. Although these differences limit the immediate relevance of our experiment to real-world cultural markets, our findings nevertheless suggest that social influence exerts an important but counterintuitive effect on cultural market formation, generating collective behavior that is reminiscent of (but not identical to) "information cascades" in sequences of individuals making binary choices (*20–22*). On the one hand, the more information participants have regarding the decisions of others, the greater agreement

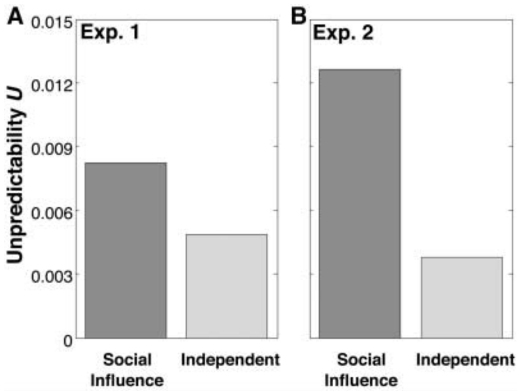

Fig. 2. Unpredictability of success for (**A**) experiment 1 and (**B**) experiment 2. In both experiments, success in the social influence condition was more unpredictable than in the independent condition. Moreover, the stronger social signal in experiment 2 leads to increased unpredictability. The measure of unpredictability u_i for a single song i is defined as the average difference in market share for that song between all pairs of realizations; i.e., $u_i = \sum_{j=1}^{W} \sum_{k=j+1}^{W} |m_{i,j} - m_{i,k}| / \binom{W}{2}$, where $m_{i,j}$ is song i's market share in world j and W is the number of worlds. The overall unpredictability measure $U = \sum_{i=1}^{S} u_i/S$ is then the average of this measure over all S songs. For the independent condition, we randomly split the single world into two subpopulations to obtain differences in market shares, and we then averaged the results over 1000 of these splits. All differences are significant ($P < 0.01$) (*18*).

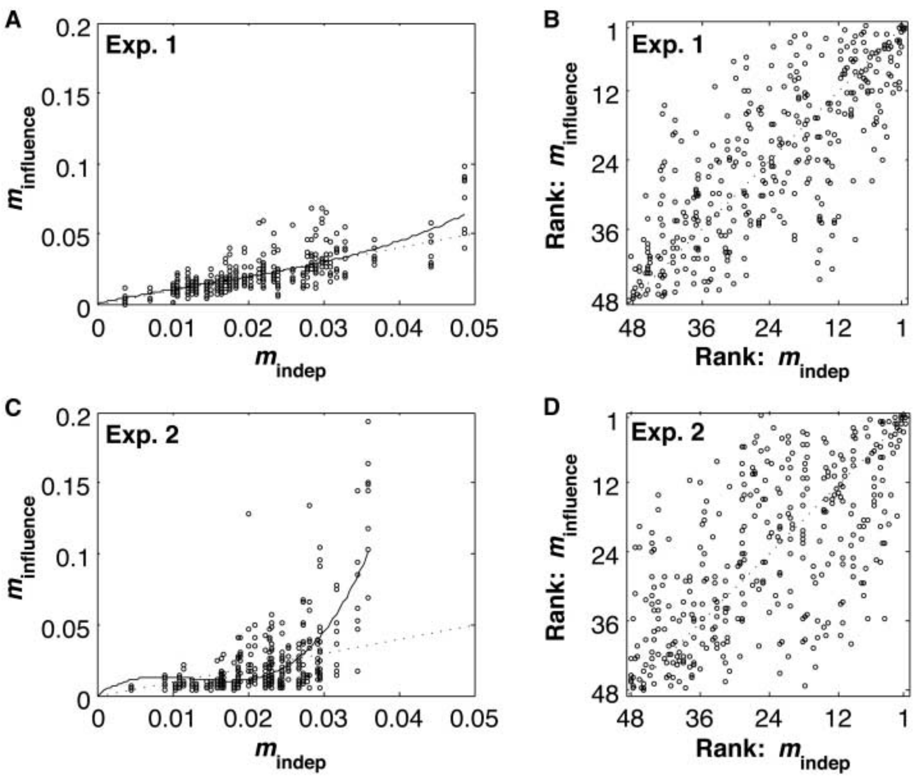

Fig. 3. Relationship between quality and success. (**A**) and (**C**) show the relationship between m_{indep}, the market share in the one independent world (i.e., quality), and $m_{\text{influence}}$, the market share in the eight social influence worlds (i.e., success). The dotted lines correspond to quality equaling success. The solid lines are third-degree polynomial fits to the data, which suggest that the relationship between quality and success has greater convexity in experiment 2 than in experiment 1. (**B**) and (**D**) present the corresponding market rank data.

they will seem to display regarding their musical preferences; thus the characteristics of success will seem predictable in retrospect. On the other hand, looking across different realizations of the same process, we see that as social influence increases (i.e., from experiment 1 to experiment 2), which particular products turn out to be regarded as good or bad becomes increasingly unpredictable, whether unpredictability is measured directly (Fig. 2) or in terms of quality (Fig. 3). We conjecture, therefore, that experts fail to predict success not because they are incompetent judges or misinformed about the preferences of others, but because when individual decisions are subject to social influence, markets do not simply aggregate pre-existing individual preferences. In such a world, there are inherent limits on the predictability of outcomes, irrespective of how much skill or information one has.

Although Web-based experiments of the kind used here are more difficult to control in some respects than are experiments conducted in physical laboratories (*18*), they have an important methodological advantage for studying collective social processes like cultural market formation. Whereas experimental psychology, for example, tends to view the individual as the relevant unit of analysis, we are explicitly interested in the relationship between individual (micro) and collective (macro) behavior; thus we need many more participants. In order to ensure that our respective worlds had reached reasonably steady states, we required over 14,000 participants—a number that can be handled easily in a Web-based experiment, but which would be impractical to accommodate in a physical laboratory. Because this "micro-macro" feature of our experiment is central to all collective social dynamics (*23*), we anticipate that Web-based experiments will become increasingly useful to the study of social processes in general.

References and Notes

1. H. L. Vogel, *Entertainment Industry Economics* (Cambridge Univ. Press, Cambridge, UK, 2004).
2. A. B. Krueger, *J. Labor Econ.* **23**, 1 (2005).
3. K. H. Chung, R. A. K. Cox, *Rev. Econ. Stat.* **76**, 771 (1994).
4. A. De Vany, *Hollywood Economics* (Routledge, London, 2004).
5. P. M. Hirsch, *Am. J. Sociology* **77**, 639 (1972).
6. W. T. Bielby, D. D. Bielby, *Am. J. Sociology* **99**, 1287 (1994).
7. R. E. Caves, *Creative Industries* (Harvard Univ. Press, Cambridge, MA, 2000).
8. R. A. Peterson, D. G. Berger, *Admin. Sci. Quart.* **16**, 97 (1971).
9. S. Rosen, *Am. Econ. Rev.* **71**, 845 (1981).
10. R. H. Frank, P. J. Cook, *The Winner-Take-All Society* (Free Press, New York, NY, 1995).
11. R. Bond, P. B. Smith, *Psychol. Bull.* **119**, 111 (1996).
12. R. B. Cialdini, N. J. Goldstein, *Annual Rev. Psych.* **55**, 591 (2004).
13. D. J. Watts, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5766 (2002).
14. P. Hedström, in *Social Mechanisms: An Analytical Approach to Social Theory*, P. Hedström, R. Swedberg, Eds. (Cambridge Univ. Press, Cambridge, UK, 1998), pp. 306–327.
15. M. Adler, *Am. Econ. Rev.* **75**, 208 (1985).
16. Available at (http://musiclab.columbia.edu).
17. Available at (http://bolt.com).
18. Materials and methods are available as supporting material on *Science* Online.
19. P. D. Allison, *Am. Sociol. Rev.* **43**, 865 (1978).
20. S. Bikhchandani, D. Hirshleifer, I. Welch, *J. Pol. Econ.* **100**, 992 (1992).
21. L. R. Anderson, C. A. Holt, *Am. Econ. Rev.* **87**, 847 (1997).
22. D. Kübler, G. Weizsäcker, *Rev. Econ. Stud.* **71**, 425 (2004).
23. J. S. Coleman, *Foundations of Social Theory* (Harvard Univ. Press, Cambridge, MA, 1990).

24. We thank P. Hausel for developing the MusicLab Web site; J. Booher-Jennings for design work; S. Hasker for helpful conversations; and A. Cohen, B. Thomas, and D. Arnold at Bolt Media for their assistance in recruiting participants. Supported in part by an NSF Graduate Research Fellowship (to M.J.S.), NSF grants SES-0094162 and SES-0339023, the McDonnell Foundation, and Legg Mason Funds.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5762/854/DC1
Materials and Methods
SOM Text
Figs. S1 to S10
Tables S1 to S4
References

6 October 2005; accepted 22 December 2005
10.1126/science.1121066

The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA

Andrew J. Barbera,[1]* Jayanth V. Chodaparambil,[2]* Brenna Kelley-Clarke,[1] Vladimir Joukov,[3] Johannes C. Walter,[4] Karolin Luger,[2] Kenneth M. Kaye[1]†

Kaposi's sarcoma–associated herpesvirus (KSHV) latency-associated nuclear antigen (LANA) mediates viral genome attachment to mitotic chromosomes. We find that N-terminal LANA docks onto chromosomes by binding nucleosomes through the folded region of histones H2A-H2B. The same LANA residues were required for both H2A-H2B binding and chromosome association. Further, LANA did not bind *Xenopus* sperm chromatin, which is deficient in H2A-H2B; chromatin binding was rescued after assembly of nucleosomes containing H2A-H2B. We also describe the 2.9-angstrom crystal structure of a nucleosome complexed with the first 23 LANA amino acids. The LANA peptide forms a hairpin that interacts exclusively with an acidic H2A-H2B region that is implicated in the formation of higher order chromatin structure. Our findings present a paradigm for how nucleosomes may serve as binding platforms for viral and cellular proteins and reveal a previously unknown mechanism for KSHV latency.

Kaposi's sarcoma–associated herpesvirus (KSHV) has an etiological role in Kaposi's sarcoma (KS), the predominant AIDS malignancy; primary effusion lymphoma (PEL); and multicentric Castleman's disease (*1*–*4*). KSHV persists as a multicopy episome in latently infected tumor cells (*5*, *6*). Viral genomes lack centromeres, which govern faithful DNA partitioning in eukaryotic cells, and use a distinct segregation mechanism in which the 1162–amino acid KSHV latency-associated nuclear antigen (LANA) tethers episomes to mitotic chromosomes. LANA is required for episome persistence, and interaction with mitotic chromosomes is essential for its function. The first 22 residues comprise the dominant LANA chromosome-association region, because the C-terminal chromosome targeting domain is unable to rescue chromosome association in mutants that are deleted for or contain specific mutations within the N-terminal region (*7*–*10*). We therefore sought to determine the chromosome docking partner of the LANA N terminus.

Genetic analysis of LANA's chromosome binding region was central to our strategy for characterization of putative docking partners. Transient assays have shown that alanine substitutions at LANA residues 5 to 7 [original amino acids were GMR (*11*)], 8 to 10 (originally LRS), or 11 to 13 (originally GRS) (termed LANA $_{5}$GMR$_{7}$, LANA $_{8}$LRS$_{10}$, and LANA $_{11}$GRS$_{13}$, respectively) (Fig. 1A) lack chromosome association, whereas LANA with alanine substitutions at amino acids 17 to 19 (originally PLT) or 20 to 22 (originally RGS) (termed LANA $_{17}$PLT$_{19}$ and LANA $_{20}$RGS$_{22}$,

[1]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. [2]Howard Hughes Medical Institute and Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523–1870, USA. [3]Department of Cancer Biology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02115, USA. [4]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: kkaye@rics.bwh.harvard.edu