Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens*

Yan Chen F. Maxwell Harper Joseph Konstan Sherry Xin Li

November 4, 2008

*We would like to thank Rachel Croson, John Duffy, Catherine Eckel, Steffen Huck, Sara Kiesler, Gueorgi Kossinets, Nancy Kotzian, Robert Kraut, Vai-Lam Mui, John List, Charles Plott, Tanya Rosenblat, Roberto Weber, and seminar participants at Community Lab (communitylab.org), Edinburgh, Innsbruck, Pittsburgh, Michigan, Yahoo! Research and ASSA 2006 (Boston, MA), ESA 2006 (Atlanta, GA) for helpful discussions, and Xinzheng Shi for excellent research assistance. The financial support from the National Science Foundation through grants no. IIS-0325837 is gratefully acknowledged. Chen: School of Information, University of Michigan, 1075 Beal Avenue, Ann Arbor, MI 48109-2112. Email: yanchen@umich.edu. Harper and Konstan: Department of Computer Science and Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455. Email: konstan@cs.umn.edu, harper@cs.umn.edu. Li: School of Economic, Political and Policy Sciences, University of Texas at Dallas, 2601 N. Floyd Road, Richardson, Texas 75083; Email: sherry.xin.li@utdallas.edu.

Abstract

We design a field experiment to explore the use of social comparison to increase contributions to an online community. We find that, after receiving *behavioral* information about the median user's total number of movie ratings, users below the median demonstrate a 530% increase in the number of monthly movie ratings, while those above the median do not necessarily decrease their ratings. When given *outcome* information about the average user's net benefit score, above-average users mainly engage in activities that help others. Our findings suggest that effective personalized social information can increase the level of public goods provision.

Keywords: social comparison, conformity, social preference, public goods, embedded online field experiment

JEL Classifications: C93, H41

1 Introduction

With the increasing popularity of the Internet, information technology is changing the way we interact, entertain, communicate and consume. In online communities, groups of people meet to share information, discuss mutual interests, play games and carry out business. Users of communities such as SourceForge (`http://sourceforge.net/`) and Wikipedia contribute information goods, which are typically shared as public goods. However, despite the popularity of online communities, many such communities fail due to nonparticipation and under-contribution. For example, Butler (2001) found that 50% of social, hobby, and work mailing lists had no traffic over a 122 day period. Under-contribution is a problem even in active and successful online communities. For example, in MovieLens (`http://www.movielens.org`), an online movie recommendation website that invites users to rate movies and, in return, makes personalized recommendations and predictions for movies the user has not already rated, under-contribution is common. More than 22% of the movies listed on the site have fewer than 40 ratings, so few that the software cannot make accurate predictions about which users would like these movies (Cosley, Ludford and Terveen 2003). Similarly, Eureka, a Xerox Corporation online information sharing system, which enables its 20,000 worldwide customer service engineers to share repair tips, also suffers from under-contribution. While many service engineers download machine repair tips from Eureka, only an estimated 20% have submitted a validated tip to the system (Bobrow and Whalen 2002). Many online communities are populated with peripheral users, who observe the community and use the contents created by others without contributing to the community content or discussions. In the P2P file sharing site Gnutella, in 2000, 25% of users share 98% of the content while 66% of users share nothing (Adar and Huberman 2000). By 2005, 85% of users share nothing (Hughes, Coulson and Walkerdine 2005). Thus, a key challenge to the online community designer is to motivate the peripheral participants to become active contributors, and the core participants to sustain and improve their contributions.

To resolve the problem of under-contribution, economists might turn to the theories of incentive-compatible mechanisms for public goods provision. However, most mechanism design theories regarding public goods rely on tax-subsidy schemes.[1] Thus, they cannot be directly applied to online communities, as these communities rely on voluntary participation and contribution of time and effort rather than monetary transfers to encourage contributions.

Furthermore, compared to traditional communities, online communities have distinct characteristics, which give the mechanism designer a new set of options. Most notably, the designer has more information than is traditionally assumed in mechanism design theory, which enables her to personalize the motivation information to each user.[2] For example, some software can track the

[1]See Groves and Ledyard (1987) for a survey of the theoretical literature and Chen (forthcoming) for a survey of the experimental literature.

[2]In dominant strategy and Nash implementations, it is usually assumed that the designer knows nothing about the

detailed activities of each user, including a user's click stream and a time stamp for each activity. From these data, the designer can infer important underlying user preferences and the time cost of each activity. Such information has been used to target customers in e-commerce, as in Amazon.com's book recommendations.[3]

In this paper, we explore how a designer might be able to motivate users to contribute contents to an online community by using personalized social information. The idea that social information might affect behavior is theorized in both social psychology and economics. In social psychology, Festinger (1954) theorizes that we compare ourselves to others who are better off for guidance, and to others who are worse off to increase our self-esteem. Since Festinger's seminal work, a large body of literature in social psychology shows that social comparisons affect behavior, since individuals gain information on what constitutes the "right behavior" in various contexts. Furthermore, social comparison theory suggests that people lean toward social comparisons in situations that are ambiguous (see Buunk and Mussweiler (2001), Suls, Martin and Wheeler (2002) for recent surveys), a condition which is true in many online communities. Although we are not aware of a mathematical formalization of social comparison theory, three special cases of this theory have been formalized in economics. In the first case, when information regarding prevalent behavior is available, people exhibit the tendency to copy this behavior, a phenomena referred to as conformity (Asch (1956), Akerlof (1980), Jones (1984), Bernheim (1994)). In the second case, when outcome information regarding other people's payoffs is available, people show distributional concerns, such as inequality aversion (Fehr and Schmidt (1999), Bolton and Ockenfels (2000)). In this case, participants in the laboratory act to reduce payoff inequalities. A third related literature model interdependent preferences, where utility functions depend not only on the absolute value of consumption, but also on either the average level of consumption (Duesenberry (1949), Pollak (1976)), or the ordinal rank in the distribution of consumption (Frank (1985), Robson (1992), Hopkins and Kornienko (2004)). Samuelson (2004)'s evolutionary model provides a justification for preferences that incorporate relative consumption effects in order to compensate for incomplete environmental information.

Most empirical studies of the impact of social information in economic decision making are conducted in the laboratory, using variants of the dictator games (e.g., Cason and Mui (1998), Krupka and Weber (2005), Duffy and Kornienko (2007)), the ultimatum bargaining games (e.g., Knez and Camerer (1995), Duffy and Feltovich (1999), Bohnet and Zeckhauser (2004)), or coor-

---

underlying distribution of preferences or the production technology, while in Bayesian implementation, it is usually assumed that the designer knows the distribution of agent preferences, but not the realization in individual agents.

[3]For example, the book *Touching the Void* (Simpson 1988), a mountain climber's account of near death in the Peruvian Andes, received good reviews and modest success when it was first published, and was soon forgotten. Years later, another mountain-climbing tragedy, *Into Thin Air* (Krakauer 1999), became a publishing sensation. Amazon began to recommend *Touching the Void* to readers who bought *Into Thin Air*. Eventually *Touching the Void* outsold *Into Thin Air* more than two to one (Anderson 2004).

dination games (Eckel and Wilson 2006). A more detailed discussion of the findings is reported in Section 3.2.

In comparison, two natural field experiments examine the effects of social information on contribution to public goods. A natural field experiment provides a bridge between a laboratory experiment and direct field observations (Harrison and List 2004). In a university fundraising campaign, Frey and Meier (2004) find that information about the average contribution in the past has a significant impact on contribution. In contrast, Shang and Croson (2005) finds that, in a public radio fundraising campaign, the most influential social information is contribution behavior drawn from the 90th to 95th percentile.

Like these two studies, we designed a natural field experiment to compare the effects of different types of social information to motivate user contributions. Our study differs from these two studies in both the type of contributions (time vs. money) and the medium of implementation (online vs. offline). We implement our experiment through a combination of email newsletters and direct modification of the MovieLens website. Furthermore, we *personalize* our social information to separately investigate the effects of social information on underperformers and overperformers, and find that they are drastically different. For designers of online communities, the effects of personalization to motivate different types of users is important, as it is technically feasible.

Additionally, this paper contributes to an emerging body of field experiment literature about eliciting participation in online communities. For example, Beenen, Ling, Wang, Chang, Frankowski, Resnick and Kraut (2004) apply the social psychology theories of social loafing and goal setting (Karau and Williams 1993) to contributions in MovieLens. They find that individuals contributed more when they were reminded of their uniqueness and when they were given specific and challenging goals. Ludford, Cosley, Frankowski and Terveen (2004) conduct a field experiment using MovieLens and test the effects of similarity and perceived uniqueness on participation in discussion groups. Dholakia, Bagozzi and Pearo (2004) use survey data to investigate the effects of group norms and social identity on participation in online communities, such as web-based chat rooms and online games. Lastly, Cosley, Frankowski, Terveen and Riedl (2007) conduct a field experiment with SuggestBot, software that recommends work to Wikipedia editors, showing that suggesting work consistent with their previous edits yields significantly more total work done.

Common to all online communities, a user's activities can have both private and pubic benefits. A user's benefit from performing an activity is called her private benefit from that activity. This private benefit includes private consumption (e.g. having a more cleanly filed and easily searchable bookmarks on del.icio.us), or the fun of playing online games. If, in addition, her activity also benefits others, i.e., it is non-rivalrous and nonexcludable, we say that this activity has public benefits. For example, in the del.icio.us community, adding bookmarks has only private benefit if a user does not reveal her bookmarks to others. She benefits from an easily searchable bookmark system. If, in addition, she makes her bookmarks public, so that others can browse or subscribe, it becomes a public good, and thus have public benefits as well (Wash and Rader 2007). We will define these terms more precisely in Section 4.

In this paper, we report a randomized field experiment on MovieLens where we send users an email newsletter which directs them to perform activities with varying degrees of private and public benefits. The newsletter contains one of two types of social information: the median number of ratings or the net benefit score of an average user in her cohort.[4]. The control group receives information about only their own past rating behavior. We then modify the interface for each user, with new shortcuts that lead to different types of contributions, including rating popular or rare movies, updating the database, inviting a buddy or just visiting MovieLens. We then track user behavior for a month after the release of the newsletter. From this experiment, we find that, after receiving *behavioral* information about the median user's total number of movie ratings, users below the median have a 530% increase in the number of monthly movie ratings, while those above the median do not necessarily decrease their ratings. When given *outcome* information about the average user's net benefit score, above-average users mainly engage in activities that help others. Our findings suggest that effective personalized social information can increase the level of public goods provision.

The rest of this paper is organized as follows. In Section 2, we introduce MovieLens. In Section 3, we present our experimental design. Section 4 presents a theoretical framework for online recommender systems and a model of social comparison. Section 5 presents the results. In Section 6, we summarize the results and discuss their implication in the design of online communities.

2 MovieLens: An Overview

MovieLens (`http://www.movielens.org`) is an online movie recommender system that invites users to rate movies and in return makes personalized recommendations and predictions for movies the user has not already rated. It is run by a research group in the Department of Computer Science and Engineering at the University of Minnesota. It is one of the most popular noncommercial movie recommender sites, and has been featured extensively by The New York Times, ABC News Nightline, and The New Yorker. Specifically, as of April 30, 2006, MovieLens has over 13 million user ratings of 9043 movies. These ratings come from just over 100,000 users, of whom approximately 15,000 were active within the past year. Since most readers are familiar with Netflix, it is important to point out the main difference between the two sites. Unlike Netflix, MovieLens does not have any DVD rental service.

To determine personalized recommendations, MovieLens uses collaborative filtering technology – an algorithmic approach to personally evaluate items for users based on the opinions of both

[4]The net benefit score is roughly the difference between the benefit a user receives from using MovieLens, and the time and effort she puts in. It is explained more precisely in Section 4

that user and the entire community of users. The underlying assumption for this technology is that those who agreed in the past tend to agree again in the future. The algorithm matches together users with similar opinions about movies, and for each user, generates a "neighborhood" of other like-minded users. Personalized recommendations for each user is generated from the ratings of these neighbors. Applications of the collaborative filtering technology include Amazon.com's book recommendation system (users who bought x also bought y), and Netflix's movie recommender system.[5] In an age of information explosion, a recommender system helps individuals find desired information. For example, in MovieLens, a user can ask MovieLens to recommend movies, either overall or within a search, and the site will return a list of movies that fulfill the user's search criteria sorted in the order of those the user is most likely to enjoy. Alternatively, the user can enter specific movies and receive a prediction of enjoyment on a 1/2- to 5-star scale. MovieLens encourages users to rate movies they have seen. Rating has two significant benefits: (a) it improves the user's profile by giving the algorithm more information about the user, and thereby may improve the quality of recommendations and predictions generated for her; and (b) it adds to the overall database of ratings, and therefore may improve the recommendations and predictions generated for others. Therefore, rating is an impure public good.

In rating movies, there are distinctions in effort and value. Movie ratings have a skewed distribution.[6] For example, the most popular movie in the system, Pulp Fiction, has been rated by nearly 50,000 users. By contrast, the bottom ten movies have zero ratings, and 75% of the movies in the system have fewer than 1100 ratings. Rating a rare movie[7] takes more work–a user needs to identify from the database one that she has seen, and most users have seen very few of them. Therefore, in the rating process, a user might need to go through many more screens of movie titles before finding one she has seen and can rate. On the other hand, rating a rare movie adds greater value to others in the community. MovieLens currently has plenty of data from which to recommend popular movies, but still needs more data to accurately and personally recommend rare ones. Furthermore, popular movies have already been seen and rated by many system users, and therefore will not be recommended to them, no matter how many new ratings they receive; in contrast, rare movies have not been seen or rated by many users (by definition) and therefore may be recommended to nearly all of the users in the system (depending on how much the system estimates those users will like them). Therefore, rating a rare movie generates higher public benefit than rating a popular movie.

[5]The recent 1 Million Dollar Netflix Prize for improving the accuracy of its movie recommendations underscores the importance of recommendation quality in online business applications. Reed Hastings, the CEO of Netflix, believes that recommendations are one of Netflix's most important advantages, especially for its non-blockbusters (Anderson 2006).

[6]The best fit distribution for the current movie ratings in the database is lognormal(2016.1, 17410), although the Kolmogorov-Smirnov test rejects it at the 5% level.

[7]In the experiment, we define a rare movie as one with fewer than 250 ratings.

The value to a user of rating a rare vs. popular movie is less clear. A user's profile is made more accurate when the user's rating: (a) differentiates the user's taste from other tastes consistent with her prior ratings, and (b) associates the user with a different set of similar users. This improved accuracy is greatest when the movie being rated has a high variance in ratings (i.e., many people like it, and many dislike it), and when that movie also has been rated by many others. Rare movies can improve a profile by associating the user with others who viewed that movie, but rarely will have as large an effect as rating a divisive popular movie.[8] For an overview of the recommender systems and the technical details of the methods, we refer the reader to Gediminas Adomavicius (2005).

In sum, rating a popular movie takes less time than rating a rare movie. Therefore, in our model developed in Section 4, we will assume that the marginal cost of rating a popular movie is less than that of rating a rare movie. Furthermore, while the private benefit from rating a popular vs. a rare movie might be comparable, the public benefit from rating a rare movie is much higher. Therefore, in the experimental design, we emphasize the private benefit from rating a popular movie, and the public benefit from rating a rare movie (Appendix A – #2 and #3).

In addition to rating movies, MovieLens users can contribute in other ways to benefit themselves or the community as a whole. For example, users can invite a buddy into the system – buddies are people who can collaborate by accessing each other's recommendations and by receiving joint recommendations (i.e., movies they each would like, and therefore might wish to see together). Adding a buddy is a good way of enhancing the user's experience (movies, and movie recommenders, are more fun with a friend). However, only 2500 MovieLens users (about 5%) have buddies in the system. Inviting a buddy is primarily valuable to the user herself, though bringing a new person to the community certainly benefits the community as a whole.

In the model in Section 4, we will assume that having a buddy contributes to the nonrating fun of using MovieLens. We also note that of the users who has a buddy, 69% has one buddy.[9] Therefore, in our model, for simplicity, we will assume that the fun of having a buddy is not determined by the number of buddies. Furthermore, we assume that finding a buddy is costless, as it required that a user knows an e-mail address to contact the person. Based on our knowledge of MovieLens, we conjecture that nearly all "buddies" are real-life acquaintances.

More recently, the movie database has been opened up to the community,[10] so users can help maintain the database by entering new movies directly into the database or by validating details of

[8] We recognize that many popular movies are not divisive, and therefore have little information-theoretic value to add to a profile.

[9] The distribution is the following: the proportion of users with one buddy is 69.36%, two buddies 15.98%, three buddies 6.33%, four and above 8.34%.

[10] Prior to 2005, the database was maintained by a single user, who did a meticulous job of database entry, but was slow in getting new movies into the database. The list of user-suggested movies to be entered into the database was so long that it became a major source of dissatisfaction among users.

existing entries (see Appendix A #5 for an example). This task provides little direct benefit to the user, but instead benefits the community as a whole. While it is possible that some users might feel that updating the database is fun or get a warm glow from the act(Andreoni 1990), based on user feedback, we view this activity as an altruistic service activity.[11]

Consequently, in our model in Section 4, we make the simplifying assumption that there is no private benefit from the act of updating the database. We assume that updating the database provides a public good to the community. Furthermore, in the experiment, we emphasize the public benefit of this activity by telling users that updating the database "will improve the quality of information in the system."

In sum, MovieLens is representative of many online communities in that the underlying collaborative filtering technology draws on user-provided information to serve each individual user and the community as a whole. The problem in such a system is how to motivate users to contribute to the (impure) public goods without using monetary incentives. This study explores the effects of social information to motivate users to contribute to the community.

3 Experimental Design

In June 2005, we launched a field study of 398 MovieLens users in order to test the effects of social information on contribution behavior. In this section, we describe our experimental design. Our experiment focuses on the impact of a personalized email newsletter sent to each of the subjects. The email newsletter contained messages that compared each subject's rating or net benefit in MovieLens with that of other users in the system. We also conducted two online surveys with our subjects before and after the experiment.

[Figure 1 about here.]

Figure 1 summarizes the experiment time line. To determine the extent to which members could understand the content of our newsletters, we conducted 14 phone interviews with MovieLens members before launching the experiment. In general, members were able to understand the information in the email newsletter. These 14 members were not included in the experiment. We refer to this phase as the Newsletter Alpha Test, which is comparable to a pilot session in a laboratory experiment.

As of 12:00:00 on June 14, 2005, there were 100,366 users in MovieLens, all of whom finished the sign-up process, which requires at least 15 movie ratings.[12] To solicit volunteers for the study, we decided to use the pool of MovieLens users who had logged in between June 2004 and June

[11] Users commented that the activity was boring. "Why can't you scrape data from IMBD?" (Cosley, Frankowski, Kiesler, Terveen and Riedl 2005).

[12] See `http://movielens.umn.edu/join`.

2005, who had rated at least 30 movies, and who had given us permission to send them email. We used the login and ratings criteria to ensure that we could calculate a user's net benefit score, which we will explain in detail in Section 4. Of all users, 5488 met our selection criteria, among whom we randomly chose 1,966 users and sent each one a recruiting email. The email contained a link to a web page containing a consent form. A total of 629 users clicked on the email link, of whom 398 consented to participate in the study.[13]

[Figure 2 about here.]

Figure 2 presents the number of users and average number of movie ratings (in brackets) of the entire population, those who met our criteria, those whom we invited, and those who participated in our study. Using two-sided Kolmogorov-Smirnov test of the equality of distributions, we find that those who met our criteria rated significantly more movies than those who did not ($p < 0.01$). Comparing those who were invited with those not invited, the number of ratings are not significantly different ($p = 0.114$). However, of those who were invited, users who voluntarily participated in the study rated significantly more movies than those who did not participate ($p < 0.01$). In sum, users in our sample rated significantly more movies compared to the population. This difference between the behavior of experiment volunteers and the population being studied is known in the experimental literature as the volunteer effects (Rosnow, Rosenthal, McConochie and Arms 1969).[14]

All study participants had the chance to earn up to three entries (by completing the two online surveys and participating in the study) in a prize drawing held at the conclusion of the study. We awarded one \$100, two \$50, and five \$20 cash prizes to participants at the end of the study. Using prize drawing is a standard method to induce users to complete online surveys (Bosnjak and Tuten 2003). In comparison, users' other activities on MovieLens, such as rating movies and inviting buddies, are part of their natural activities on the site, which we do not need to influence with a prize. We collect user behavioral data during the month before the recruiting email was sent out (weeks 1-4 in Figure 1) when behavior had not been influenced by any experimental stimulus, and after the personalized newsletter was sent out (weeks 7-10 in Figure 1), leaving out the recruiting and pre-survey period (weeks 5 and 6).

[13]Based on the post-experiment survey of the participants, 75% are male, 91% have at least college education, and 76% are between age 20 and 40.

[14]In general, we are likely to see volunteer effects in an online community study for two reasons beyond those experienced in offline studies. First, frequent users are more likely to encounter advertisements or recruitment messages. We avoided this challenge by recruiting through email, a mechanism through which each user receives a single invitation whether that user visits hourly or infrequently. Second, loyal users are more likely to feel affinity for the community and therefore might be more likely to participate in studies conducted on the site. We recognize this limitation, but randomly assigned users among conditions so that all conditions would be similarly biased. We recognize that volunteer recruitment may limit our generalizability to those similar to the volunteers, but argue this limitation is true in most similar experiments.

3.1 Pre-Experiment Survey

Users who consented to participate in this study were immediately redirected to an online 10-question survey. The first purpose of this survey was to elicit users' perceptions of their benefits and costs from using MovieLens, using questions drawn from our earlier study of online recommender systems (Harper, Li, Chen and Konstan 2005). We used these survey responses in combination with information on participants' historical usage of MovieLens to compute net benefit scores for those in the Net Benefit treatment. The second purpose of this survey was to discover how users believed they compared with other users in the study, in terms of how many movies they rated and their net benefit from using the system. 383 of the 398 subjects in the experiment completed this survey. A copy of the pre-experiment survey is posted at
`http://www.grouplens.org/data/mlsurvey2005/index.html`.

3.2 Personalized Email Newsletter and Modified MovieLens Interface

Approximately two weeks after sending the initial invitation to participate in the study, we sent a personalized email newsletter to each subject. We randomly divided the 398 subjects into the three experimental groups. A user's experimental group determined the type of email newsletter the user would receive in the study. The first treatment group, Rating Info, received a newsletter indicating how many movies they had rated compared with the median user in their group. The second treatment group, Net Benefit, received a newsletter indicating how much net benefit they obtained from using MovieLens compared with other users. Finally, the Control group received a newsletter with only information about their own ratings profile.[15]

Findings from social psychology have suggested that people are more responsive to comparisons with people sharing similar related attributes than to comparisons with dissimilar others (Suls, Martin and Wheeler 2002). In our study, we hoped to avoid comparing a new user with users who had been using the system for years. Thus, we further subdivided the Rating Info and the Net Benefit groups into three membership cohorts, New, Mid and Old, based on a user's date of registration with MovieLens. Table 1 presents the characteristics of each of the three membership cohorts. Although we did not divide the control group into cohorts in the experiment, whenever a treatment group is compared to the control in the analysis, we compare the corresponding membership cohorts respectively. In the two treatments, there are approximately equal number of users in each cohort. The numbers in brackets are the number of active users who rated movies, updated the database or invited a buddy during the two-month period of data collection, i.e., the months before the recruiting email and after the newsletter was sent out.

[15]The exception to the random assignment of users to experimental groups is the 15 users who did not complete the pre-experiment survey. They were assigned to the Rating Info and the Control groups, as we did not have the information to compute their net benefit score. In subsequent analyses, we include all 398 users. We repeat all analyses excluding the 15 users who did not complete the pre-survey and find that the main results still hold.

[Table 1 about here.]

All three newsletters are similar in design. Each is formatted in html, although users with text-only email clients received a text-only version.[16] Each design contained a header, with the MovieLens logo, and some statistics about the number of MovieLens members, movies, and ratings. Below the header, there were three sections. The first section contained personalized information according to the subject's experimental group, as described below. The second section contained a short news item about recent feature additions to MovieLens. The final section was a reminder about the research study prizes. Sample email newsletters are included in Appendix A (#1 - #3).

The first section of the newsletter, which contained personalized information about the subject, was the source of our experimental manipulation. While all three experimental groups received different types of personalized information, all of the newsletters contained the same five links: (1) rate popular movies, (2) rate rare movies, (3) invite a buddy to use MovieLens, (4) help us update the MovieLens database, and (5) just visit the MovieLens home page. These links were clarified by neighboring text that explained the effect of these actions on a subject's own as well as others' experience in MovieLens. For example, the link "rate rare movies" was followed by the text "rating rare movies will help others get more movie recommendations." While all contained the same links, the links were grouped differently according to the experimental condition. Furthermore, depending on the participant's experimental group, the email contained one of these additional messages.

Subjects in the Rating Info treatment received a message about how many movies they had rated compared with other users in their cohort. Their newsletter contained the following text:

> "Ever wondered how many movies you've rated compared with other users like you? You have rated [ ] movies. Compared with other users who joined MovieLens around the same time as you, you've rated [more, fewer, about as many] movies than the median (the median number of ratings is [ ]).

Two main options followed this text, randomly ordered. One main option was to rate more movies, followed by the links to rate popular movies and to rate rare movies. The other main option was to try new MovieLens features. Under this heading we provided two links, one to invite a buddy to use MovieLens and another to help maintain the MovieLens database, again randomly ordered. Below these links was the link to the MovieLens home page.

Participants in the Net Benefit treatment received a message emphasizing their net benefits from using MovieLens compared with the net benefits of other users. Their newsletter contained the following text:

[16] Each was sent in dual format, html and text-only. The email client of the user automatically chose which one to display.

> "We have calculated the net benefit that you get from MovieLens, a measure of the enjoyment and the value you receive minus the time and effort you put in. Your net benefit score is [ ]. Compared with other users who joined MovieLens around the same time as you, your net benefit is [above, below, about] average (the average net benefit score is [ ])."

In a footnote in the email newsletter, we explain the concept of net benefit: "The net benefit score is a measure of the total benefit you receive from using MovieLens minus the time and effort you put in. The total benefit you receive includes the value of movie recommendations you get from MovieLens, and your enjoyment from rating movies and other fun activities, such as browsing movies. This score is computed by using a mathematical model constructed in one of our earlier studies. The information used includes your activities on MovieLens and your responses to related questions in the survey. The score ranges from 60 to 90." This score is calculated based on Equation (1) in Section 4. The benefit to a user maps into the private benefit a user receives from the system.

We again provided two main options, randomly ordered. One main option was to "increase your net benefit score," followed by the links to invite a buddy to use MovieLens and to rate popular movies, randomly ordered. The other main option was to "help others increase their net benefit scores," followed by links to help maintain the MovieLens database and to rate rare movies, again randomly ordered. Below these links was the link to the MovieLens home page.

An important design decision is the type of social information provided in the experiment. In other studies of social comparison, different social information has been selected and presented to the participants. Several studies present the decision(s) of *one other* participant and find mixed results. Cason and Mui (1998) find that, in sequential dictator games, although observation of behavior of one other participant constraints subjects from moving towards self-regarding choices, the effect is modest as behavior of one randomly chosen other might not change individual beliefs about what constitutes the appropriate behavior. Duffy and Feltovich (1999) find that observation of behavior of one randomly chosen pair influences behavior in different ways in the repeated ultimatum and best-shot games. In a coordination game in Eckel and Wilson (2006), observation of the move of one player affects behavior of other players only when this player has high status. In comparison, in the public radio fundraising field experiment, Shang and Croson (2005) find that the most influential social information is contribution behavior of a donor drawn from the 90th to 95th percentile of previous contributions, although participants do not know the percentile of the comparison target. A second type of social information is the *complete ranking* of all participants, such as in Duffy and Kornienko (2007), who find that such ranking information has significant effects on giving in dictator games, however, it might not be applicable to a large population such as that in our experiment. Finally, Bohnet and Zeckhauser (2004) present the *average* offer in ultimatum bargaining games and find that this information activates the social norm of equal split.

In a university fundraising field experiment, Frey and Meier (2004) also present information about the average contribution behavior of the student population in the past and find significant impact on contribution. In a closely related study of binary dictator games, Krupka and Weber (2005) let subjects observe the decisions of four players from previous experiments and find a significant jump in sharing when the number of the other players who share increases from two to three, consistent with the effect of a social norm.

Based on findings in other studies and the public goods nature of our experiment, we choose the median or average as the social information presented to our participants. Note that, in the Rating Info treatment, we use the *median* rating as the social information rather than the mean, as the distribution of the number of ratings is right skewed due to the presence of power users. Using the median rating rather than the average rating ensures comparable sample sizes across above-, about, and below-median groups and across membership cohorts. More importantly, information about the median allows users to infer the behaviors of the numerical majority used in conformity theory. In contrast, in the Net Benefit treatment, we use the *average* net benefit score, as the distribution of the net benefit scores is symmetrically centered. As a result, the medians and the averages are almost the same across the three membership cohorts of participants. Based on the results of our alpha test, most of MovieLens users understood the concept of median, and had intuitive knowledge about how to interpret net benefit scores. All of them understood the comparison of their standing relative to that of their cohorts.

Finally, the Control group received a message about their participation in MovieLens without any comparison to other users. Their newsletter contained the following text:

> "Here are some statistics about your ratings behavior for one popular movie genre. About [ ] of the movies that you've rated are comedies. Your average rating in this genre is [ ]."

This message was followed by the same five links and explanations offered to the Rating Info and Net Benefit treatments, although the links were not grouped. The order of the first four links was randomized, with the link to visit the MovieLens home page at the bottom.

Comparing the treatments and control newsletters, we note that they differ in more than the social information dimension. For example, in the Net Benefit treatment, the activity links are proceeded by "To increase your net benefit score" or "To help others increase their net benefit score," while in the control, they are proceeded by "Interested in getting more out of MovieLens? Here are some options." These different wordings were crafted to make the newsletter look and feel natural to participants of our field experiment. However, we do not rule out the possibility that they might fundamentally change how people interpret these options. Consequently, the comparison between a treatment and the control might be affected by this confound. In contrast, the comparisons of users in the same treatment, e.g., above and below average users in the Net Benefit treatment, are immune from these potential confounds.

Subjects who visited MovieLens, either by clicking on the newsletter's links or otherwise, were given a slightly modified interface with the four links from the email newsletter included in the "shortcuts" pane of the main MovieLens interface - visible from each page in the system (Appendix A #4). These four links behaved exactly as they did in the email, but were logged differently so that we could differentiate between the different types of actions. Following shortcut conventions at MovieLens, the links on the site were not annotated with explanatory information.

3.3 Post-Experiment Survey

We waited for one month after we sent the email newsletter to give the subjects a chance to use the system. At the conclusion of the month, we emailed the users again with an invitation to take a second survey. This survey included MovieLens related questions, questions modified from the General Social Survey, the Big Five personality survey,[17] and questions on demographics. 310 of the subjects (78%) completed this survey. A copy of the survey is posted at `http://www.grouplens.org/data/mlsurvey2005/index.html`.

4 A Theoretical Framework

In this section, we first set up a static model of online recommender systems, which extends the one developed in Harper et al. (2005) by incorporating new MovieLens features. We then extend the static model into a two-period model which incorporates social comparisons based on our experimental design. The theoretical model produces a set of hypotheses for experiment.

4.1 A Static Model

We first outline a static model in the neoclassical framework with self-interested agents. This model is appropriate for an online community where social information has been largely unavailable before the implementation of our experiment. The MovieLens community is entirely virtual and nearly anonymous. Until recently, users were not made aware of the presence of others, except through their limited understanding of the recommendation process. For most users, this recommendation system is a tool that helps them keep track of, find, and recommend movies.[18] Therefore, absent of social information, a neoclassical model captures the basic features and motivations in the MovieLens community.

[17]The Big Five measures five broad dimensions of personality (Goldberg 1993). It is now among the most widely accepted and used models of personality.

[18] Since the experiment described in this paper, a social tagging system has been added to the site, which increases the opportunity for social visibility.

In our model, there are n users. Let X_i be the total number of ratings from user i, and $X_i = X_i^p + X_i^r$, where X_i^p and X_i^r are the number of popular and rare movies[19] user i has rated respectively. Let d_i be the number of movie entries updated by user i. Let $d = \sum_{i=1}^{n} d_i$ be the total number of validated movie entries in the database, which is a public good.

Based on survey data (Harper et al. 2005), a user's benefit from using MovieLens comes from three sources. The most important benefit is the quality of the movie recommendations, $Q_i(X_i, \sum_{j\neq i} X_j)$, which depends on one's own ratings that the algorithm uses to infer a user's taste, and the stock of ratings in the system. Based on the characteristics of the algorithm, we assume that $Q_i(\cdot,\cdot)$ is concave in both its components, i.e., more ratings from a user increase the quality of her recommendations, but at a decreasing rate. More total ratings by others in the system also increase the quality of recommendations, at a decreasing rate. For analytical tractability, we assume that $Q_i(\cdot,\cdot)$ is additively separable. We denote the marginal benefit from the quality of recommendations as γ_i. The second source of benefit comes from rating fun, $f_i(X_i)$, as identified by the enjoyment derived from rating movies and voicing opinions. We assume that $f'(\cdot) > 0$, and $f''(\cdot) \leq 0$. Finally, users may also enjoy non-rating activities, h_i, including enjoyment from browsing and having a buddy. As explained in Section 2, for simplicity, we assume that non-rating fun is independent of the number of webpages browsed (for simplicity), or the number of buddies, as 69% of the users have only one buddy. We also assume that finding a buddy is costless, because to invite a buddy involves clicking a link, and filling out the name and email address of the person, then clicking a "Submit" button. As we opened up the database for the experiment, we add a fourth component of benefit derived from a validated database, $v_i(d)$, where $v_i(\cdot)$ is concave and twice continuously differentiable.

In our model, we further assume that there is a cost associated with rating movies. The (total) cost function of rating movies, $c_i(X_i)$, measures the amount of time that agent i needs to rate X_i movies. Assume $c_i(X_i)$ is convex, i.e., the marginal cost is positive, $c_i^{'}(X_i) > 0$, and $c_i^{''}(X_i) \geq 0$ for all $i \in N$. This assumption captures the feature that the marginal cost of rating either remains constant or increases with the number of ratings. A distinction between popular and rare movies is that the marginal cost of rating a popular movie is less than that of rating a rare movie, i.e., $dc_i/dX_i^p < dc_i/dX_i^r$. Similarly, we assume that the cost of updating the database is $c_i^d(d_i)$, where $c_i^d(\cdot)$ is also convex.[20]

Taking into consideration all benefits and costs of using MovieLens, we specify a user's neoclassical utility function as

[19] Recall that, in the experiment, we define a rare movie as one with fewer than 250 ratings.

[20] Based on the time stamp of activities in our experimental logfiles, we find that rating a popular movie takes a median user 9 seconds (based on 537 movie rating events), while rating a rare movie takes a median user 11 seconds (based on 30 movie rating events). Note that the latter might be an underestimate of the actual time cost because of the small sample size. Updating a database entry, however, takes a median user 90 seconds (based on 348 events).

$$\pi_i(X_i, \sum_{j \neq i} X_j) = \gamma_i Q_i(X_i, \sum_{j \neq i} X_j) + f_i(X_i) + h_i + v_i(d) - c_i(X_i) - c_i^d(d_i). \quad (1)$$

We assume additive separability to get a close-form solution for our empirical analysis (Harper et al. 2005). In our experiment, we use Equation (1) to compute a user's net benefit score from using MovieLens.[21]

In what follows, we extend the static neoclassical model to a two-period model which incorporates the two different kinds of social information in our experiment treatments, and derive theoretical predictions for the experiment.

4.2 Behavioral Comparison: Rating Info Treatment

We first extend the model to incorporate the effect of social information on behavior. Recall, in the Rating Info treatment, we give each participant information about her own number of movie ratings and the number of ratings by the median user in her membership cohort. Based on the social comparison theory, and conformity theory in particular, we expect that this information will have an effect on user behavior.

Mathematical models of conformity either directly assume disutility from non-conforming behavior (Akerlof 1980) or derive equilibrium behavior from a signalling model (Bernheim 1994) where users care about their "intrinsic" utility as well as their status. In a pooling equilibrium, when status is sufficiently important, individuals with heterogeneous preferences conform to a homogeneous standard of behavior. In this subsection, we extend Akerlof's (1980) reduced form model to characterize the effect of behavioral comparison with the median user on individual behavior.

In this model, the basic unit of time is one month. Suppose the newsletter is released at the end of month t. After the release, users have information to compare themselves with the median user in their cohort. Let x_i^τ be user i's total number of ratings in month τ. Then $X_i^t = \sum_{\tau=1}^t x_i^\tau$ is the total number of ratings from user i up to time t. Let X_m^t be the total number of ratings from the median user at time t. We analyze the behavioral data in the month following the release of the newsletter, x_i^{t+1}, and compare this data to that in the month before, x_i^t. With a slight abuse of notation, we use π_i^t to denote the net benefit in period t.

A user's utility function after learning the median user's rating information can be expressed as follows,

$$u_i(X_i^{t+1}, \sum_{j \neq i} X_j^{t+1}, X_m^{t+1}) = \pi_i^{t+1} - g_i(|X_i^{t+1} - X_m^{t+1}|), \quad (2)$$

[21] We set $d_i = 0$ when computing the net benefit score prior to the start of the experiment, and used the number of database entries during the month of the experiment when computing the score at the end of the experiment.

where

$$\pi_i^{t+1} = \gamma_i Q_i(X_i^{t+1}, \sum_{j \neq i} X_j^{t+1}) + f_i(x_i^{t+1}) + h_i + v_i(d^{t+1}) - c_i(x_i^{t+1}) - c_i^d(d_i^{t+1}), \tag{3}$$

and where $g_i(\cdot)$ captures the disutility from deviating from the social norm. We assume that $g_i(\cdot) \geq 0$, for $i \neq m$, indicating that a user is either indifferent or suffers disutility from deviating from the social norm. We further assume that this disutility weakly increases with greater deviation from the norm, i.e., $g_i'(\cdot) \geq 0$. While Equation (2) might not be the most general functional form which captures the effects of social comparison, it maps into our experimental design the best. In subsequent discussions, we index a user below the median in the number of ratings as l, and one above the median as h.

Observation 1. Comparing rating behavior in the month before and after the release of the newsletter, we have the following results:
(a) The median user's behavior remains the same, i.e., $x_m^{t+1} = x_m^t$, or $\Delta x_m = 0$;
(b) Users below the median will rate more movies in the month after compared to the month before, i.e., $x_l^{t+1} \geq x_l^t$, or $\Delta x_l \geq 0$;
(c) Users above the median will rate fewer movies in the month after compared to the month before, i.e., $x_h^{t+1} \leq x_h^t$, or $\Delta x_h \leq 0$; and
(d) Users in the control group will rate the same number of movies in the month after compared to the month before, i.e., $x_c^{t+1} = x_c^t$, or $\Delta x_c = 0$;

Proof: See Appendix B.

Observation 1 compares each group's rating behavior in the month after the newsletter with its behavior in the month before. Theory predicts that users from both ends of the spectrum will change their rating behaviors. In our theoretical framework, users in the control group do not receive any social information about ratings, so their rating behavior remains the same. However, in reality, there might be spurious events not captured in our model which can cause the rating behavior of users to change. An analysis method to address this issue is to compare the difference in behavior in the treatment with that in the control groups. This Observation provides a theoretical benchmark for such analysis in Section 5. Observation 1 is a common prediction that can be made by several theories, including the informational social comparison theory (Samuelson 2004), conformity (Akerlof (1980), Bernheim (1994)), anchoring and priming, and mimicking. We present it in the context of MovieLens for completeness and for stating testable hypotheses for our experiment.

In the following proposition, we compare the groups within the Rating Info treatment with each other.

Proposition 1. *When conforming to the new social norm is sufficiently important, i.e., when $g_i'(\cdot)$ is sufficiently large,*
(a) Users below the median will rate at least as many movies as the median user in the month after receiving the newsletter, or $x_l^{t+1} \geq x_m^{t+1}$;
(b) Users above the median will rate at most as many movies as the median user in the month after receiving the newsletter, or $x_h^{t+1} \leq x_m^{t+1}$.
(c) At the aggregate level, we should observe conformity to the median, $|X_i^{t+1} - X_m^{t+1}| \leq |X_i^t - X_m^t|$.

Proof: See Appendix B.

Proposition 1 indicates that, if conforming to the social norm is sufficiently important, the distance between a user's total number of ratings and the total number of ratings of the median user at time $t+1$ is no greater than the distance at time t when the newsletter was released. In other words, we expect the distribution to be tighter after the release of the median rating information. Together, Observation 1 and Proposition 1 provide a theoretical benchmark for the data analysis of our Rating Info treatment.

4.3 Outcome Comparison: Net Benefit Treatment

In contrast to the Rating Info treatment, where the information regarding a median user's *behavior* is presented, in the Net Benefit treatment, we present the *outcome* information, i.e., the user's own net benefit score and that of the average user. In the social psychology literature on social comparison, people compare themselves to others in both the behavior and outcome dimensions (e.g., Suls et al. (2002), Lockwood and Kunda 1997). We now extend the model developed in Subsection 4.2 to the comparison in outcomes, i.e., the net benefit scores. Rewriting Equation 2 in the outcome space, we get

$$u_i^{t+1} = \pi_i^{t+1} - g_i(|\pi_i^{t+1} - \pi_a^{t+1}|). \tag{4}$$

Again, $g_i(\cdot)$ captures the disutility from deviating from the social norm, i.e., the average net benefit score, with the same assumptions on the properties of $g_i(\cdot)$. To avoid excessive notation, we use a, l and h to index users with net benefit scores about, below and above average, respectively. We first look at an average user, i.e., $\pi_a \doteq \bar{\pi}$. Equation 4 implies that she maximizes her neoclassical utility function,

$$u_a^{t+1} = \pi_a^{t+1}. \tag{5}$$

For a user with a net benefit score below average, her utility function is:

$$u_l^{t+1} = \pi_l^{t+1} - g_l(\pi_a^{t+1} - \pi_l^{t+1}), \tag{6}$$

Therefore, when she is below average, she suffers disutility which is increasing in the distance between her net benefit and the average user's net benefit.

For a user with a net benefit score above average, her utility function is:

$$u_h^{t+1} = \pi_h^{t+1} - g_h(\pi_h^{t+1} - \pi_a^{t+1}), \tag{7}$$

Therefore, when she is above average, she again suffers disutility which is increasing in the distance between her net benefit and the average user's net benefit.

In a special case when $g_i(\cdot)$ is linear, our model has the same functional form as the inequality aversion model of Fehr and Schmidt (1999).[22] In this case, the coefficient, g_l in Equation (6), can be interpreted as the degree to which user l envies the average user, while the coefficient, g_h in Equation (7), can be interpreted as the degree of a user's charity concerns. While the Fehr and Schmidt (1999) model assumes that $g_l \geq g_h$ and $g_h \in [0, 1)$, we do not impose these additional assumptions (cite Dirk Engelmann's work on the robustness of these assumptions). The following proposition characterizes user response to the outcome information when $g_i(\cdot)$ is linear.

Proposition 2. *For the Net Benefit treatment, we expect the following results:*
(a) For an average or a below-average user, it is a dominant strategy to rate popular movies, and a dominated strategy to rate rare movies or to update the database.
(b) For an above-average user, there exists a $g_h^* \in (0, 1)$, such that

- *when $g_h < g_h^*$, it is a dominant strategy to rate popular movies, and a dominated strategy to rate rare movies or to update the database;*
- *when $g_h \geq g_h^*$, it is a dominant strategy to rate rare movies and to update the database, and a dominated strategy to rate popular movies.*

Proof: See Appendix B.

Proposition 2 predicts that an average or a below-average user is more likely to rate popular movies than to rate rare movies or to update the database. For an above-average user, if she has competitive preferences ($g_h < 0$) or is sufficiently selfish ($g_h < g_h^*$), she is more likely to rate popular movies than to rate rare movies or to update the database. However, if she is sufficiently charitable ($g_h \geq g_i^*$), she is more likely to choose activities which benefit the community, i.e., rating rare movies or updating the database.

Proposition 2 enables us to compare behaviors across groups. If the fraction of users with sufficient charity concerns is positive, we expect that the above-average users will be more likely to rate rare movies or to update the database compared to the average or below-average users or

[22] We do not attempt to review the large literature on social preference here. Rather, we refer the reader to several key models in this literature, including Rabin (1993), Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Falk and Fischbacher (2006), Cox, Friedman and Gjerstad (2007).

those in the control group. Similarly, we expect that the average or below-average users are more likely to rate popular movies than the above average group. Finally, we expect that the average users will behave similarly to the control group.

5 Results

In this section, we present our data analysis and main results. After tracking user behavior in the month after receiving the email newsletter, we find significant and interesting behavioral responses to the social information we presented in the newsletter.

There are some common features that apply throughout our analysis. First, since the median user's behavior can be idiosyncratic, in the analysis, we compare the rating behavior of the below- and above-median groups with that of the median group,[23] rather than the median user. Similarly, in the Net Benefit treatment, we compare the above- and below-average users with that of the average group, rather than the average user. Second, we note that the Invite-a-Buddy shortcut did not attract the attention of our users.[24] There were a total of seven buddies invited for the entire subject pool, too small for any meaningful statistical comparisons across treatments. Therefore, in reporting the results, we focus on movies ratings and database updating. Lastly, since 275 out of 398 participants (see Table 1) were active in the two-month period, we report separate results for all users vs. active users.

We first verify that the pre-experiment distributions of total movie ratings between each of the treatment groups and the control group come from the same distribution. The results of Kolmogorov-Smirnov tests cannot reject the equality of distribution functions except for the comparison of old users between the Net Benefit treatment and the control group.[25]

[Figure 3 about here.]

Figure 3 presents an overview of user rating behavior in the Rating Info treatment and control groups, comparing the month before (the white bar) and the month after the newsletter (the black bar). The left panel includes all users, while the right one includes only active users. Compared to the month before, the effects of social information on post-newsletter behavior are striking. For the Rating Info group, users below the median have a 530% increase in the total number of movie ratings, while those above the median decrease their monthly ratings by 62%. Movements from

[23] The median group is defined as the 1/6 of users with lifetime ratings above and below the median, i.e. the middle 1/3 of the users for each membership cohort. It is kept constant over time.

[24] We speculate that this might be due to the demographics of our subject pool. Based on the post-experiment survey, more than 70% of our subjects are male between the age of twenty and forty.

[25] P-values from the Kolmogorov-Smirnov tests between the Rating Info treatment and the control groups are 0.99 (overall), 0.84 (New), 0.97 (Mid) and 0.85 (Old). P-values of the same tests between the Net Benefit treatment and the control groups are 0.62 (overall), 1.0 (New), 0.98 (Mid) and 0.02 (Old).

both ends converge towards the median, although the effect of social information is more dramatic for those below the median. In comparison, the about median group has a 290% increase in the number of ratings in the month after compared to the month before, which is not predicted by conformity theory. However, a closer examination of the about median group reveals that most of the increase comes from those who are *actually* below the median (88% for new users, 91% for mid users and 79% for old users), which is consistent with conformity theory.

The striking change in post-newsletter behavior might be attributed to the social information, or to any spurious trends absent of the social information, including a regression to the mean effect. To differentiate the two effects, we compare the change in behavior in the Rating Info treatment and the control group. If the change in behavior in the Rating Info treatment is due to a regression to the mean effect, we expect to observe it in the control group as well. Specifically, we compute the difference in the number of movie ratings in the month before and after the release of the newsletter, $\Delta x_{E,i} = x_{E,i}^{t+1} - x_{E,i}^{t}$, for each experimental treatment $E \in \{$R(ating Info), C(ontrol)$\}$, and for the below-, about- and above-median groups $i \in \{l, m, h\}$. We then check whether there are significant differences between the corresponding treatment and the control groups. Recall that the control group was never divided into the below, about and above median subgroups in the experiment itself. This division is only used in the analysis to investigate any regression to the mean effect. If the change in behavior is due to the social information, based on Observation 1, we expect that, compared to the corresponding subgroups in the control, the change in movie ratings will be larger for the below-median group, about the same for users in the median group, and smaller for users in the above-median group.

[Table 2 about here.]

Result 1 (Rating Info vs. Control)**.** Compared to the control group, the change in movie ratings within the Rating Info group is significantly larger for the below-median and the median groups, and about the same for users in the above-median groups.

Support. Table 2 presents the average difference in the total number of ratings for each group in the Rating Info treatment and control groups, with differential effects on the new, mid and old users. Using the Wilcoxon rank-sum test, we reject the null hypothesis $\Delta x_{R,l} = \Delta x_{C,l}$ in favor of $\Delta x_{R,l} > \Delta x_{C,l}$ ($p = 0.01$ for all users, $p = 0.00$ for active users). Furthermore, we reject the null $\Delta x_{R,m} = \Delta x_{C,m}$ in favor of $\Delta x_{R,m} > \Delta x_{C,m}$ ($p = 0.05$ for all users, $p = 0.02$ for active users). However, we fail to reject the null $\Delta x_{R,h} = \Delta x_{C,h}$ in favor of $\Delta x_{R,h} < \Delta x_{C,h}$ ($p = 0.41$ for all users, $p = 0.24$ for active users). ∎

Result 1 confirms that the social information in the Rating Info treatment group has a significant effect on behavior in the below-median and the median groups. Compared to Observation 1, only the prediction for the below-median group is confirmed. While the outcome for the median group is different from the theoretical prediction, as we noted before, more than 80% of the increase in the median group comes from users who are *actually* below the median. We now proceed to analyze behavioral changes within the Rating Info treatment.

[Table 3 about here.]

Result 2 (Conformity in ratings)**.** In the month after the release of the newsletter, among active users in the Rating Info treatment, those below the median rate significantly more movies than their median counterparts. Among all users in the Rating Info treatment group, the above-median users rate significantly more movies than the median users. At the aggregate level, the average distance between an active user and her median counterpart's number of ratings is significantly smaller in the month after than that in the month before.

Support. Table 3 presents our hypotheses and the corresponding Wilcoxon rank sum test statistics. The alternative hypotheses are derived from Proposition 1 in Section 4. Among active users (lower panel), below-median users rate more movies than median users ($p = 0.02$ overall and $p = 0.01$ for new users). Among all users (upper panel), $x_h^{t+1} = x_m^{t+1}$ cannot be rejected in favor of $x_h^{t+1} < x_m^{t+1}$ ($p = 0.97$ overall and $p = 0.98$ for new users). However, we can reject $x_h^{t+1} = x_m^{t+1}$ in favor of $x_h^{t+1} > x_m^{t+1}$ ($p = 0.03$ overall and $p = 0.02$ for new users). To test part (c) of Proposition 1, we compute the distance between each user i and her median counterpart's cumulative ratings in the month before the newsletter, $|X_i^t - X_m^t|$, and the month after, $|X_i^{t+1} - X_m^{t+1}|$. Pair-sample t-tests reject the null of equal distance in favor of the alternative hypothesis that the distance in the month after (347.9) is smaller than that in the month before (354.5) for the active users ($n = 99$, $p = 0.02$). If we include all users in the Rating Info treatment, however, we fail to reject the null in favor of the alternative ($n = 134$, $p = 0.42$). ∎

Since Proposition 1 holds when conforming to the social norm is sufficiently important, i.e., when $g_i'(\cdot)$ is sufficiently large, we expect the results to be different for users with different tendencies to conform. To investigate this difference, we partition the users into *conforming* and *non-conforming* types based on their post-experiment survey responses to a pair of questions designed to measure the tendency to conform, i.e., "I will stick to my opinion if I think I am right, even if others disagree," and its reversed version, "I will change the opinion I express as a result of an onslaught of criticism, even though I really don't change the way I feel." The correlation of the responses to these two questions is 0.5 ($p < 0.01$). We categorize a user as a conforming type if she responds "agree" or "strongly agree" to the first question, and "disagree" or "strongly disagree" to the second, and a non-conforming type otherwise. Out of the 83 users who answered both two questions, 58 are categorized as conforming and 25 non-conforming types.[26]

[26] As a robustness check, we construct two alternative measures of conforming tendency based on the two questions separately, and find that the results are consistent with those based on this composite measure.

Before presenting results combining the behavioral and survey data, we mention several common caveats associated with the post-experiment survey data. First, only 78% of the users responded to the survey. Second, 71% of the users answered every question. Lastly, among users who answered a set of questions and their reversed versions, the answers are not always consistent.[27] Therefore, one should use caution interpreting the survey results.

[Table 4 about here.]

Table 4 presents our two-type analysis in two panels. Panel A, which corresponds to parts (a) and (b) of Proposition 1, presents the comparison of the post-experiment monthly ratings between the below-median, median, and above-median users. Results indicate that the prediction of Proposition 1 (a) is only significant for the conforming types ($p = 0.03$), while the prediction of Proposition 1 (b) is not significant for either type ($p > 0.10$).

Panel B, which corresponds to Proposition 1 (c), presents the average distance between each user i and her median counterpart's cumulative ratings in the month before, $|X_i^t - X_m^t|$, and the month after the newsletter, $|X_i^{t+1} - X_m^{t+1}|$, for the conforming and non-conforming types respectively. We find that, for the conforming types, the distance from median in the month after is significantly smaller than that in the month before ($p = 0.04$, one-side paired-sample t-tests). For the non-conforming types, although the average distance also reduces, the reduction is not statistically significant ($p = 0.20$). Results therefore support Proposition 1 (c).

In sum, while Proposition 1 predicts the behavior of users below the median well, its prediction does not hold for users above the median, who rate significantly more movies than the median users. Furthermore, the cohorts most responsive to the median rating information are the new users, who might be more malleable.

Both Results 1 and 2 suggest that the median rating information has a more dramatic effect on the below-median group (a 530% increase in total ratings compared to the month before) compared to the above-median group (a 62% decrease in total ratings). We speculate that this disparity in effect might be due to an interaction between conformity and competitive preferences. In MovieLens, the system exhorts the users to rate more movies. For example, in the new user tour, one screen says "Remember: the more movies you rate, the more accurate MovieLens' predictions will be." Therefore, rating more movies might be perceived as a socially desirable course of action, which could, in turn, trigger competitive preferences, i.e., more ratings are better. For the below-median group, conformity and competitiveness work in the same direction, whereas for the above-median users, conformity theory predicts a decrease in the number of monthly ratings, while competitive preference predicts an increase. User responses to the post-experiment survey are consistent with this speculation.

[27] For example, for the two questions on conformity, consistent answers would yield a correlation of 1.0, rather than 0.5, which might reflect different interpretations of the same questions.

[Figure 4 about here.]

Figure 4 presents the change in ratings (Δx_i) for the below-, about- and above-median groups as a function of self-reported competitiveness in the survey.[28] The average number of ratings by below-, median and above-median users is represented by white, grey and black bars, respectively. While below-median users for all competitiveness levels increase their number of ratings, the more competitive users increase their number of ratings by a larger amount. By contrast, for above-median users, the change in ratings is negatively correlated with their competitiveness. Specifically, noncompetitive users have the largest decrease in the number of ratings, followed by the neutral group, while the competitive users have a slight increase in their number of ratings. Median users follow the same pattern, with the exception of the competitive users in the group.

Recall that, to keep the experimental treatments and the control strategically comparable, all users in the experiment are provided with the same five shortcuts. While conformity theory predicts that the number of ratings moves towards the median, it does not predict any systematic pattern for how users might differ in the number of database entries updated. Indeed, we find that users below-, about- and above-median are not significantly different in the number of database entries they provide. Comparing the Rating Info treatment group with the control group, we find that users in the control group provide weakly significantly more entries in the database ($p = 0.09$, one-tailed Wilcoxon rank sum test). One plausible explanation is that updating the database is a relatively new feature in MovieLens and the novelty of this feature might have attracted the attention of the users in the control group, since they do not receive any social information.

In sum, in the Rating Info treatment group, social information significantly changes user rating behavior. By reporting the median user's rating in each relevant MovieLens membership cohort, we observe a shift of behavior towards the median from both ends of the spectrum. The effect is more dramatic for the below-median users than for the above-median users. For both groups, however, we observe an interaction between conformity and competitive preferences. For below-median users, more competitive users have larger increases in the number of ratings, whereas for above-median users, more competitive users have a smaller decrease in the number of ratings.

In the Net Benefit treatment group, we provide net benefit information to investigate whether we can leverage users' distributional preferences to contribute to high-cost public goods such as rating rare movies or updating the database. We now examine the results for this group.

[Figure 5 about here.]

Figure 5 presents an overview of user behavior in the Net Benefit treatment, comparing behavior in the month before (the white bar) and the month after (the black bar) the newsletter. The left

[28]In the post-experiment survey, participants were asked to indicate their agreement on a scale of 1 (strongly disagree) to 5 (strongly agree) with the following statement, "It's achievement, rather than popularity with others, that gets you ahead nowadays." They are considered to have a noncompetitive preference if they choose 1 or 2, a neutral preference if they chose 3, and a competitive preference otherwise.

column presents the behavior of all users, while the right column presents that of the active users. Since updating the database was not available prior to the experiment, the last row does not contain any white bars.

We first verify that behavioral changes in the treatment group are due to user responses to the social information in the newsletters by comparing changes in behavior in the treatment and control groups. Since updating the database was not available prior to the experiment, we examine changes in popular and rare movie ratings compared to the respective behaviors in the control group.

[Table 5 about here.]

Result 3 (Net Benefit vs. Control)**.** The increases in popular movie ratings for the below-average and the average groups are both significantly greater than the control group.

Support. Table 5 presents the average difference in the number of popular movie ratings for each group in the Net Benefit treatment and control groups. The increase in popular movie ratings is significantly greater for the below-average group than for the control group ($p = 0.02$ for mid users among all users, and 0.07 for active users, one-sided Wilcoxon rank sum tests). Furthermore, the increase in popular movie ratings for the average users is also significantly greater than that in the control group ($p < 0.01$ for all and active users, one-sided Wilcoxon rank sum tests). ∎

Result 3 indicates that the change in popular movie rating in the Net Benefit group is indeed caused by the social information in the newsletter. We conduct similar analysis for the rare movie ratings. However, as there are fewer rare movies rated, we cannot reject the hypothesis that below-average, average, and above-average groups are the same as the respective control groups ($p = 0.72, 0.57$ and 0.51 respectively, two-sided Wilcoxon rank sum tests). Therefore, compared to the control group, the social information provided induces the below-average and average users to rate more popular, but not more rare movies, i.e., among the rating options, they prefer the more selfish to the more other-regarding one.

We next compare the behavior of different groups within the Net Benefit treatment group in the month after the newsletter. We examine three activities: the number of popular movies rated, the number of rare movies rated, and the number of database entries updated. We summarize the main findings in Result 4.

Result 4 (Inequality Aversion)**.** In the month after receiving the newsletter, users receiving different net benefit information have significantly different activity levels:
(a) Popular movie ratings: The average users rate significantly more popular movies than those below or above average;
(b) Rare movie ratings: The above-average users rate significantly more rare movies than those below-average;
(c) Database entries: The above-average users contribute 94% of the new updates in the database

from the Net Benefit treatment group, significantly more than the average or the below-average users.

Support. All p-values presented are from Wilcoxon rank sum tests:
(a) Popular movie ratings: $x_a^{p,t+1} = x_l^{p,t+1}$ is rejected in favor of $x_a^{p,t+1} > x_l^{p,t+1}$ at $p = 0.03$ (all users). Likewise, $x_a^{p,t+1} = x_h^{p,t+1}$ is rejected in favor of $x_a^{p,t+1} > x_h^{p,t+1}$ at $p = 0.03$ (active users).
(b) Rare movie ratings: $x_h^{r,t+1} = x_l^{r,t+1}$ is rejected in favor of $x_h^{r,t+1} > x_l^{r,t+1}$ at $p = 0.01$ (all users).
(c) Database entries: $d_h^{t+1} = d_l^{t+1}$ is rejected in favor of $d_h^{t+1} > d_l^{t+1}$ at $p < 0.01$ (all users), $p = 0.01$ (active users). Likewise, $d_h^{t+1} = d_a^{t+1}$ is rejected in favor of $d_h^{t+1} > d_a^{t+1}$ at $p < 0.01$ (all users), $p = 0.01$ (active users). ∎

Result 4 indicates that, overall, users with above average net benefit scores mainly engage in activities that raise the net benefit of others, i.e., rating rare movies and updating the database. It does not, however, separate users with high and low philanthropic concerns. Since Part (b) of Proposition 2 predicts that above-average users with sufficiently high philanthropic concerns will rate rare movies and update the database, we now use survey data to classify users by their altruism score.

We construct an altruism score from subject responses to six of the questions designed to measure altruism in the post-experiment survey, where a higher score represents a greater self-reported altruistic preference.[29]

[Table 6 about here.]

Table 6 reports the above-average user activities, their altruism score (Low, Middle, High), hypotheses based on part (b) of Proposition 2, and the p-values for one-sided t-tests. The results indicate that (1) the average popular movie ratings decrease with the altruism score; (2) the average rare movie ratings increase with the altruism score; and (3) the average database entries increase with the altruism score. These results are largely consistent with the theoretical predictions. Among these three activities, however, only rare movie rating comparisons between those with low and high ($p = 0.010$ for all and $p = 0.024$ for active above-average users), and those with

[29]Participants were asked to indicate their level of agreement with the following statements regarding their personalities, "I see myself as someone who a) is helpful and unselfish with others; b) can be cold and aloof; c) is considerate and kind to almost everyone; d) likes to cooperate with others; e) is often on bad terms with others; f) feels little concern for others; g) is on good terms with nearly everyone." (For statements a), c), d) and g), we code the answers "strongly agree," "agree," "neutral," "disagree" and "strongly disagree" as 2, 1, 0, -1, and -2, respectively. For statements b), e), and f), we code the answers "strongly agree," "agree," "neutral," "disagree" and "strongly disagree" as -2, -1, 0, 1 and 2, respectively. Summing each individual's responses across the above questions yields a score that ranges from -5 to 13 with a mean of 4 and standard deviation of 3.8. We bin the scores into three categories, where category 1 (Low) includes those who are more than one standard deviation below the mean, category 2 (Middle) includes those within one standard deviation of the mean, and category 3 (High) includes those who are more than one standard deviation above the mean.

middle and high ($p = 0.037$ for all and $p = 0.066$ for active above-average users) altruism scores are statistically significant.

[Table 7 about here.]

Summarizing all treatments, Table 7 presents eight Tobit specifications of contribution behavior using social information categories as well as demographic characteristics as independent variables. The dependent variables include the number of movie ratings in the Rating Info treatment (specifications 1 and 2), the number of popular (3 and 4) and rare movie ratings (5 and 6), and the number of database entries (7 and 8) in the Net Benefit treatment in the month after the newsletter. The independent variables include Pre-rating (the number of movies rated in the month before), Above (users with lifetime rating above the median or net benefit score above the mean), Below, MovieLens Age, Male, Age,[30] Education (years of education), and occupation variables including Student, CompMath (Computer and Math occupations), EdTLib (Education, Training and Library occupations), and a constant. Omitted variables include Median, Female, and other occupations. In specifications (1) and (2), we find that the below-median users rated significantly more movies compared to the median group, consistent with Result 2. Unlike in Result 2, however, the comparison between the above-median and the median users is not significant, which might be due to the smaller number of observations in the regression analysis, or the inclusion of demographic variables. Among the demographic variables, MovieLens Age is negatively correlated with the number of movies ratings in the month after, consistent with our previous result that new users are most responsive to the social information.

In the Net Benefit treatment, we find that, compared to the average users, the above-average users rated significantly less popular movies (4), significantly more database entries (7 and 8), while the below-average users rated significantly less rare movies (5 and 6), which is largely consistent with Result 4. While gender has no effect on behavior, older users rated significantly more rare movies. Lastly, users with more education finished significantly less database entries.

Throughout the analysis, we have assumed that the social information provided in the newsletter causes the changes in user behavior. A competing hypothesis suggests that anchoring might have caused the changes. To investigate this possibility, we use our control data, where no social information was provided. Instead, each user in the control group is provided with their personal statistics. If anchoring is the mechanism at work, we expect that the personal statistics in the control would bias the behavior in the month after. We find that the correlation between the information provided in the newsletter, i.e., the anchor, in each treatment and the number of movies rated in the month after, is small and statistically insignificant at the 10% level in both treatments and the control group, which indicates that it is unlikely that changes in behavior is primarily caused by anchoring.

[30] Age is a categorical variable in the post experiment survey. The users are given several categories: below 15, 15 - 19, 20-29, 30-39, 40-49, 50 and above.

Lastly, for both the Rating Info and Net Benefit treatments, we compare the distribution of rankings in the month before and after to check whether there are any changes in the distribution.[31] More specifically, we are interested in whether the significant changes in the amount of movie ratings and database updating have moved some below-median (or below-average) users to above the median (or average) in movie ratings (or net benefit scores), and vice versa. Using the Spearman and Kendall rank correlation tests, we find that the correlation of rankings for movie ratings is close to one in the Rating Info treatment, whereas the correlation of the net benefit scores is strongly positive. Furthermore, we can reject the null hypothesis of ranking independence at the 1% significance level for all tests. Therefore, the relative ranking of users remain largely unchanged despite a significant amount of work by various groups of users during the month after the newsletter, an effect known as the Red Queen Effect, taken from Lewis Carroll's (1871) *Through the Looking-Glass*, where the Red Queen said, "Now *here*, you see, it takes all the running you can do, to keep in the same place."

At the aggregate level, although the total number of ratings in the Rating Info treatment does not change from the month before (2569) to the month after (2556) the newsletter, we do observe a 530% increase in the below-median group. For the Net Benefit treatment, while the number of monthly movie ratings has a 59% increase from 1216 in the month before to 1928 in the month after, above-average users rate more rare movies and contribute 94% of the new updates in the database, activities that mostly benefit others. In contrast, the control group has a 33% decrease in the number of movies ratings (from 2431 to 1632), however, users in the control group contribute 273 new updates in the database. In our entire subject pool, the monthly ratings have a 1.6% decrease from before (6216) to after the intervention (6116), while there is a net increase of 417 new updates in the database.

From a mechanism designer's perspective, to increase the overall contribution to online communities, it is important to personalize the social information, which has disparate effects on different groups of people. For example, the median rating information is effective to increase ratings for users with a low number of ratings, but not for those with a high number of ratings. In comparison, the average net benefit score can motivate users with above-average scores to increase the level of costly activities which mainly help others, and those with below- and about-average scores to increase levels of activities which mostly benefit themselves. Personalization is feasible and low-cost, especially for online communities.

6 Conclusion

The Internet enables the formation of online communities and collaboration on a scale never seen before. Many popular websites, such as Wikipedia, MySpace and YouTube, are based entirely on

[31] We thank John Duffy for suggesting this part of the analysis.

content contributed by their members. The challenge facing designers and managers of such online communities is to motivate members to sustain and improve their contributions.

In this study, we investigate the impact of social comparisons as a natural, non-pecuniary incentive mechanism for motivating contributions to an online community. Specifically, we use email newsletters to let members of an online movie recommender community know how they compare with other members in terms of movie ratings and net benefits. We find that, after receiving *behavioral* information about the median user's total number of movie ratings, users below the median show a 530% increase in the number of monthly movies ratings, while those above the median decrease their monthly ratings by 62%. Furthermore, we find that the effects of social comparisons are most dramatic for the below-median users, consistent with an interaction between conformity and competitive preferences. Additionally, we find that when given *outcome* information about the average user's net benefit score from the system, the average users rate significantly more popular movies, while users with net benefit scores above average contribute 94% of the new updates in the database, consistent with social preference theory.

Our findings have significant implications for both the mechanism designers and managers of online communities. We demonstrate that social information has significant effect on user contribution to public goods. From the perspective of designers and managers of an online community, our findings indicate that one can effectively classify users and personalize their messages to increase the amount of high-value work done by members of an online community. For example, in the case of MovieLens, for users with a low number of ratings, information on the median user's ratings can induce significantly more ratings. For users with high net benefit scores, information on their scores and those of an average user can trigger their distributional concerns and lead to an increase in contributions to the database updating and rating of rare movies. What is particular intriguing is that average users, upon learning that they are about average, can be challenged to increase their ratings as well.

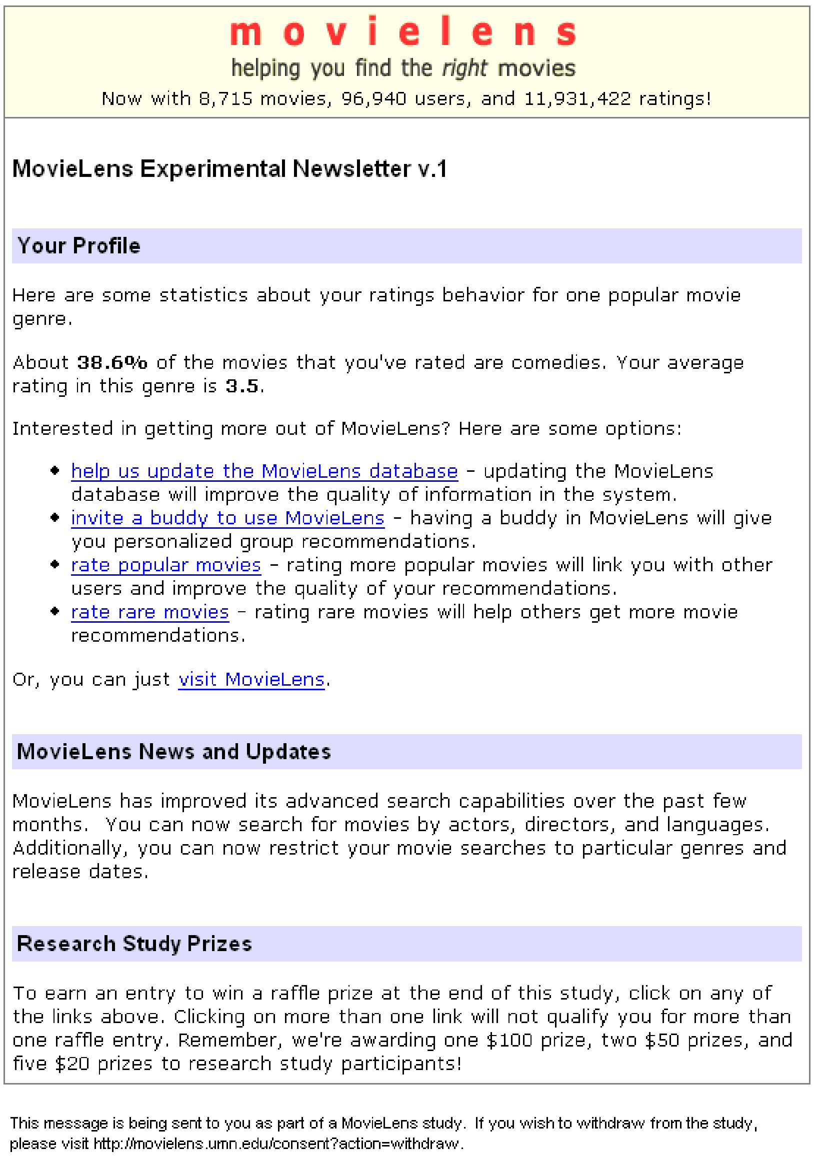
Our findings also contribute to the theoretical literature on conformity and social norms. Most existing models have the characteristic that agents suffer disutility when they deviate from the social norm (e.g., Akerlof (1980), Bernheim (1994)). Our results indicate that an interaction between conformity and competition is an important factor which has been ignored. When the social norm, such as movie ratings, contributes to the common good, conformity works in the same direction as competition for people below the median, whereas they work in opposite directions for those above the median, resulting in a more dramatic effect on the low end of the spectrum than on the high end.

In sum, our results indicate that social comparison can provide an effective non-pecuniary incentive to motivate contributions to online communities. One limitation of this study is that MovieLens is largely a leisure community. It would be interesting to examine whether we can replicate our results in work-oriented online communities. To explore this possibility, we are conducting projects on online reference communities, such as Google Answers. Furthermore, in our study, we investigate social comparisons with peers, through information provided about the median or average user. In practice, we also observe other forms of social comparisons, such as leaderboards in the ESP game (`http://www.espgame.org/`), and contribution-based status levels at Slashdot (`http://slashdot.org/`). In future work, we hope to study different forms of social comparisons and evaluate their effects on user behavior and the growth of online public information goods.

APPENDIX A. Screen Shots

In this appendix, we include one example newsletter for each treatment. Other newsletters have the same format and layout, except for the individual specific numbers and comparison phrases.

1. Email Newsletter: Control Group



movielens

helping you find the *right* movies

Now with 8,715 movies, 96,940 users, and 11,931,422 ratings!

MovieLens Experimental Newsletter v.1

Your Profile

Here are some statistics about your ratings behavior for one popular movie genre.

About **38.6%** of the movies that you've rated are comedies. Your average rating in this genre is **3.5**.

Interested in getting more out of MovieLens? Here are some options:

- help us update the MovieLens database - updating the MovieLens database will improve the quality of information in the system.
- invite a buddy to use MovieLens - having a buddy in MovieLens will give you personalized group recommendations.
- rate popular movies - rating more popular movies will link you with other users and improve the quality of your recommendations.
- rate rare movies - rating rare movies will help others get more movie recommendations.

Or, you can just visit MovieLens.

MovieLens News and Updates

MovieLens has improved its advanced search capabilities over the past few months. You can now search for movies by actors, directors, and languages. Additionally, you can now restrict your movie searches to particular genres and release dates.

Research Study Prizes

To earn an entry to win a raffle prize at the end of this study, click on any of the links above. Clicking on more than one link will not qualify you for more than one raffle entry. Remember, we're awarding one $100 prize, two $50 prizes, and five $20 prizes to research study participants!

This message is being sent to you as part of a MovieLens study. If you wish to withdraw from the study, please visit http://movielens.umn.edu/consent?action=withdraw.

2. Email Newsletter: Rating Info Treatment (Below Median)

movielens

helping you find the *right* movies

Now with 8,715 movies, 96,940 users, and 11,931,422 ratings!

MovieLens Experimental Newsletter v.1

Your Profile

Ever wondered how many movies you've rated compared with other users like you?

You have rated **287** movies. Compared with other users who joined MovieLens around the same time as you, **you've rated fewer movies** than the median (the median number of ratings is 500).

If you'd like to **rate more movies**, here are some options:

- rate popular movies - rating more popular movies will link you with other users and improve the quality of your recommendations.
- rate rare movies - rating rare movies will help others get more movie recommendations.

If you'd like to **try new features**, you may want to:

- invite a buddy to use MovieLens - having a buddy in MovieLens will give you personalized group recommendations.
- help us update the MovieLens database - updating the MovieLens database will improve the quality of information in the system.

Or, you can just visit MovieLens.

MovieLens News and Updates

MovieLens has improved its advanced search capabilities over the past few months. You can now search for movies by actors, directors, and languages. Additionally, you can now restrict your movie searches to particular genres and release dates.

Research Study Prizes

To earn an entry to win a raffle prize at the end of this study, click on any of the links above. Clicking on more than one link will not qualify you for more than one raffle entry. Remember, we're awarding one $100 prize, two $50 prizes, and five $20 prizes to research study participants!

This message is being sent to you as part of a MovieLens study. If you wish to withdraw from the study, please visit http://movielens.umn.edu/consent?action=withdraw.

3. Email Newsletter: Net Benefit Treatment (Below Average)

movielens

helping you find the *right* movies

Now with 8,715 movies, 96,940 users, and 11,931,422 ratings!

MovieLens Experimental Newsletter v.1

Your Profile

We have calculated the *net benefit** that you get from MovieLens, a measure of the enjoyment and the value you receive minus the time and effort you put in.

Your net benefit score is **61**. Compared with other users who joined MovieLens around the same time as you, your net benefit is **below average** (the average net benefit score is 66).

To **increase your net benefit score**, you may want to:

- invite a buddy to use MovieLens - having a buddy in MovieLens will give you personalized group recommendations.
- rate popular movies - rating more popular movies will link you with other users and improve the quality of your recommendations.

To **help others increase their net benefit scores**, you may want to:

- help us update the MovieLens database - updating the MovieLens database will improve the quality of information in the system.
- rate rare movies - rating rare movies will help others get more movie recommendations.

Or, you can just visit MovieLens.

MovieLens News and Updates

MovieLens has improved its advanced search capabilities over the past few months. You can now search for movies by actors, directors, and languages. Additionally, you can now restrict your movie searches to particular genres and release dates.

Research Study Prizes

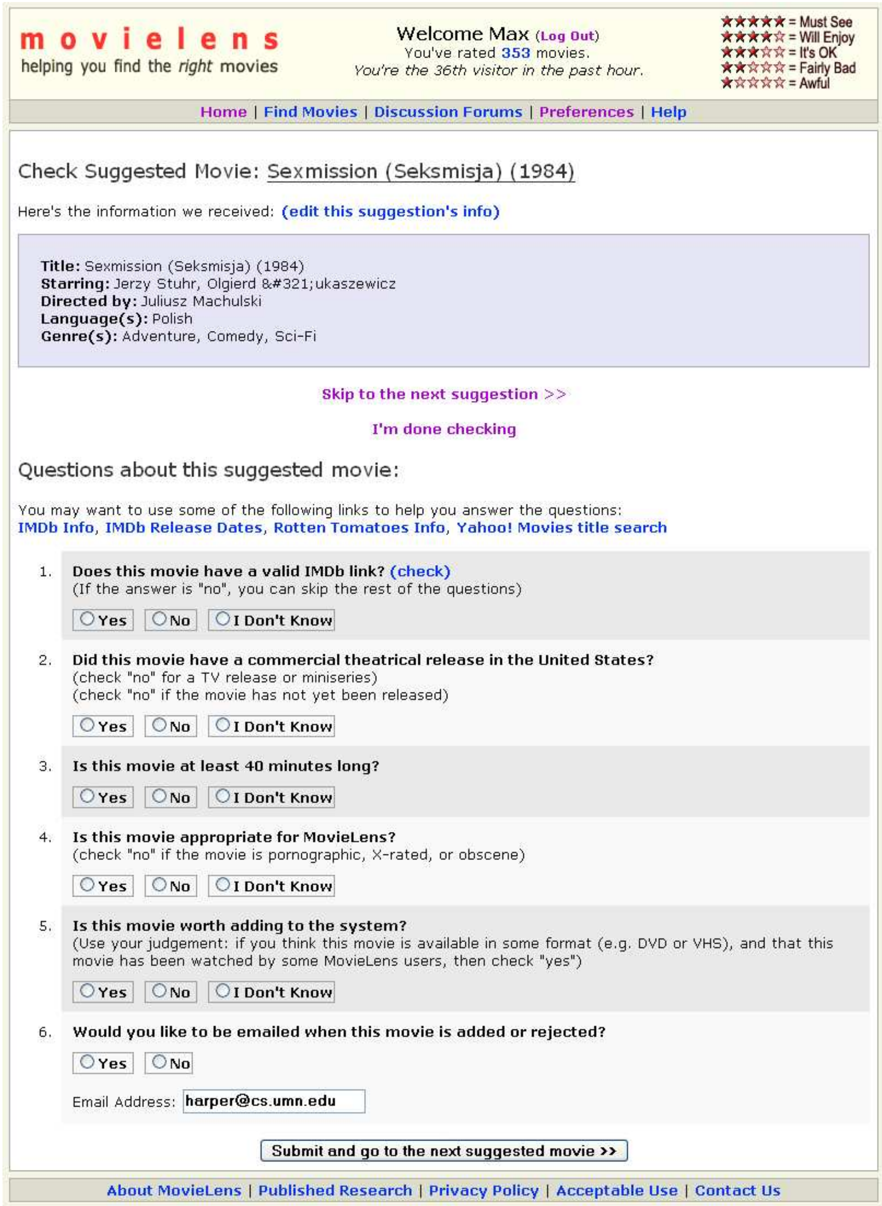
To earn an entry to win a raffle prize at the end of this study, click on any of the links above. Clicking on more than one link will not qualify you for more than one raffle entry. Remember, we're awarding one $100 prize, two $50 prizes, and five $20 prizes to research study participants!

*The net benefit score is a measure of the total benefit you receive from using MovieLens minus the time and effort you put in. The total benefit you receive includes the value of movie recommendations you get from MovieLens, and your enjoyment from rating movies and other fun activities, such as browsing movies. This score is computed by using a mathematical model constructed in one of our earlier studies. The information used includes your activities on MovieLens and your responses to related questions in the survey. The score ranges from 60 to 90.

This message is being sent to you as part of a MovieLens study. If you wish to withdraw from the study, please visit http://movielens.umn.edu/consent?action=withdraw.

4. Modified MovieLens Interface: Shortcuts

5. Updating the Database



movielens
helping you find the *right* movies

Welcome Max (Log Out)
You've rated 353 movies.
You're the 36th visitor in the past hour.

★★★★★ = Must See
★★★★☆ = Will Enjoy
★★★☆☆ = It's OK
★★☆☆☆ = Fairly Bad
★☆☆☆☆ = Awful

Home | Find Movies | Discussion Forums | Preferences | Help

Check Suggested Movie: Sexmission (Seksmisja) (1984)

Here's the information we received: **(edit this suggestion's info)**

Title: Sexmission (Seksmisja) (1984)
Starring: Jerzy Stuhr, Olgierd Łukaszewicz
Directed by: Juliusz Machulski
Language(s): Polish
Genre(s): Adventure, Comedy, Sci-Fi

Skip to the next suggestion >>

I'm done checking

Questions about this suggested movie:

You may want to use some of the following links to help you answer the questions:
IMDb Info, IMDb Release Dates, Rotten Tomatoes Info, Yahoo! Movies title search

1. **Does this movie have a valid IMDb link? (check)**
(If the answer is "no", you can skip the rest of the questions)
Yes | No | I Don't Know

2. **Did this movie have a commercial theatrical release in the United States?**
(check "no" for a TV release or miniseries)
(check "no" if the movie has not yet been released)
Yes | No | I Don't Know

3. **Is this movie at least 40 minutes long?**
Yes | No | I Don't Know

4. **Is this movie appropriate for MovieLens?**
(check "no" if the movie is pornographic, X-rated, or obscene)
Yes | No | I Don't Know

5. **Is this movie worth adding to the system?**
(Use your judgement: if you think this movie is available in some format (e.g. DVD or VHS), and that this movie has been watched by some MovieLens users, then check "yes")
Yes | No | I Don't Know

6. **Would you like to be emailed when this movie is added or rejected?**
Yes | No

Email Address: harper@cs.umn.edu

Submit and go to the next suggested movie >>

About MovieLens | Published Research | Privacy Policy | Acceptable Use | Contact Us

APPENDIX B

Proof of Observation 1: We analyze the three types of users separately.

(a) For the median user, $i = m$, at time t, she solves

$$\max_{x_m^t} \pi_m^t = \gamma_m Q_m(X_m^t, \sum_{j \neq m} X_j^t) + f_m(x_m^t) + h_m + v_m(d^t) - c_m(x_m^t),$$

which yields the following first order condition,

$$\gamma_m \frac{\partial Q_m}{\partial X_m^t} + f_m' - c_m' = 0. \tag{8}$$

Let x_m^t be the solution to Equation (8). At time $t+1$, we assume that the median user believes that she continues to be the median, therefore, $g_m(\cdot) = 0$. Thus she solves

$$\max_{x_m^{t+1}, d_m^{t+1}} \pi_m^{t+1} = \gamma_m Q_m(X_m^{t+1}, \sum_{j \neq m} X_j^{t+1}) + f_m(x_m^{t+1}) + h_m + v_m(d^{t+1}) - c_m(x_m^{t+1}) - c_m^d(d_m^{t+1}),$$

which yields the following first-order conditions,

$$\gamma_m \frac{\partial Q_m}{\partial X_m^{t+1}} + f_m' - c_m' = 0, \tag{9}$$

$$v_m' - c_m^{d'} = 0. \tag{10}$$

Let $\{x_m^{t+1}, d_m^{t+1}\}$ be the solution to Equations (9) and (10). Comparing Equations (8) and (9), it immediately follows that the median user's rating behavior should remain the same before and after the newsletter, i.e., $x_m^{t+1} = x_m^t$.

(b) For any user below the median, i.e., $l \neq m$ and $X_l^t < X_m^t$, at time t, she solves

$$\max_{x_l^t} \pi_l^t = \gamma_l Q_l(X_l^t, \sum_{j \neq l} X_j^t) + f_l(x_l^t) + h_l + v_l(d^t) - c_l(x_l^t),$$

which yields the following first order condition,

$$\gamma_l \frac{\partial Q_l}{\partial X_l^t} + f_l' - c_l' = 0. \tag{11}$$

Let x_l^t be the solution to Equation (11). At time $t+1$, she solves

$$\max_{x_l^{t+1}, d_l^{t+1}} \pi_l^{t+1} - g_l(X_m^{t+1} - X_l^{t+1}),$$

which yields the following first-order conditions,

$$\gamma_l \frac{\partial Q_l}{\partial X_l^{t+1}} + f_l' - c_l' + g_l' = 0, \tag{12}$$

$$v_l' - c_l^{d'} = 0. \tag{13}$$

Let $\{x_l^{t+1}, d_l^{t+1}\}$ be the solution to Equations (12) and (13). Since π_l is concave in x_l^{t+1} and $g_l' \geq 0$, it follows from Equations (11) and (12) that $x_l^{t+1} \geq x_l^t$. That is, a user who is below the median will increase her monthly ratings after receiving the newsletter.

(c) For any user above the median, i.e., $h \neq m$ and $X_h^t > X_m^t$, at time t, she solves

$$\max_{x_h^t} \pi_h^t = \gamma_h Q_h(X_h^t, \sum_{j \neq h} X_j^t) + f_h(x_h^t) + h_h + v_h(d^t) - c_h(x_h^t),$$

which yields the following first order condition,

$$\gamma_h \frac{\partial Q_h}{\partial X_h^t} + f_h' - c_h' = 0. \tag{14}$$

Let x_h^t be the solution to Equation (11). At time $t+1$, she solves

$$\max_{x_h^{t+1}, d_h^{t+1}} \pi_h^{t+1} - g_h(X_h^{t+1} - X_m^{t+1}),$$

which yields the following first-order conditions,

$$\gamma_h \frac{\partial Q_h}{\partial X_h^{t+1}} + f_h' - c_h' - g_h' = 0, \tag{15}$$

$$v_h' - c_h^{d'} = 0. \tag{16}$$

Let $\{x_h^{t+1}, d_h^{t+1}\}$ be the solution to Equations (15) and (16). Since π_h is concave in x_h^{t+1} and $g_h' \geq 0$, it follows from Equations (14) and (15) that $x_h^{t+1} \leq x_h^t$. That is, a user who is above the median will decrease her monthly ratings after receiving the newsletter.

(d) The analysis of users in the control group is the same as that for the median group, as they do not receive any social information. Therefore, $g_c(\cdot) = 0$, and $x_c^{t+1} = x_c^t$.

∎

Proof of Proposition 1: If conforming to the social norm is sufficiently important, i.e., if g_i' is sufficiently large, Equation (12) implies that a user below the median will rate more movies in the month after the newsletter than the median user, i.e., $x_l^{t+1} \geq x_m^{t+1}$. Similarly, (15) implies that a user above the median will rate fewer movies in the month after the newsletter than the median user, i.e., $x_h^{t+1} \leq x_m^{t+1}$. Since $|X_i^{t+1} - X_m^{t+1}| = |X_i^t - X_m^t + x_i^{t+1} - x_m^{t+1}|$, it follows that

$$|X_i^{t+1} - X_m^{t+1}| \leq |X_i^t - X_m^t|. \tag{17}$$

Equation (17) shows that the distance between a user's total number of ratings and those of the median user at time $t+1$ is no greater than the distance at time t when the newsletter was released.
∎

Proof of Proposition 2:

(a) For the average user, $i = a$, she maximizes $u_a^{t+1} = \pi_a^{t+1}$. In the newsletter, we inform the user that rating popular movies will increase her own net benefit score (π_a^{t+1}), while rating rare movies or updating the database will help others increase their net benefit score. Therefore, for an average user, rating popular movies dominates rating rare movies or updating the database.

(b) For a below-average user, l, her utility function is $u_l^{t+1} = (1 + g_l)\pi_l^{t+1} - g_l\pi_a^{t+1}$. Since rating popular movies will increase her own net benefit score, π_l^{t+1}, while rating rare movies or updating the database will help others increase their net benefit score, which increases π_a^{t+1}, rating popular movies dominates rating rare movies or updating the database.

(c) For a user with a net benefit score above average, h, her utility function is $u_h^{t+1} = (1-g_h)\pi_h^{t+1} + g_h\pi_a^{t+1}$. We discuss several cases.

- $g_h \leq 0$: for a competitive or selfish user, rating popular movies improves her own net benefit score, π_h^{t+1}, and therefore, dominates rating rare movies or updating the database.
- $g_h = 1$: for a selfless user, rating rare movies or updating the database improves others' net benefit scores, π_a^{t+1}, and therefore, dominate rating popular movies.
- $g_h \in (0, 1)$: there exists a g_h^* such that
  - when $g < g_h^*$, it is a dominant strategy to rate popular movies (and a dominated strategy to rate rare movies or to update the database).
  - When $g \geq g_h^*$, it is a dominant strategy to rate rare movies or to update the database (and a dominated strategy to rate popular movies). ∎

References

Adar, Eytan and Bernardo A. Huberman, "Free Riding on Gnutella," *First Monday*, October 2000, *5* (10).

Adomavicius, Alexander Tuzhilin Gediminas, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering*, June 2005, *17* (6), 734–749.

Akerlof, George A., "A Theory of Social Custom, of Which Unemployment May be One Consequence," *The Quarterly Journal of Economics*, June 1980, *94* (4), 749–775.

Anderson, Chris, "The Long Tail," *Wired Magazine*, October 2004, *12* (10), 170–177.

__ , *The Long Tail: Why the Future of Business Is Selling Less of More*, New York, NY: Hyperion, 2006.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, June 1990, *100* (401), 464 – 477.

Asch, Solomon E., "Studies of Independence and Conformity. A Minority of One Against a Unanimous Majority," *Psychological Monographs*, 1956, *70* (416), 1–70.

Beenen, Gerard, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, and Robert E. Kraut, "Using Social Psychology to Motivate Contributions to Online Communities," in "Proceedings of ACM CSCW 2004 Conference on Computer Supported Cooperative Work," Vol. Chicago, IL. 2004.

Bernheim, Douglas, "A Theory of Conformity," *Journal of Political Economy*, October 1994, *102* (5), 841–877.

Bobrow, Daniel G. and Jack Whalen, "Community Knowledge Sharing in Practice: The Eureka Story," *Reflections, Journal of the Society for Organizational Learning*, Winter 2002, *4* (2), 47–59.

Bohnet, Iris and Richard J. Zeckhauser, "Social Comparison in Ultimatum Bargaining," *Scandinavian Journal of Economics*, September 2004, *106* (3), 495–510.

Bolton, Gary E. and Axel Ockenfels, "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, March 2000, *90* (1), 166–193.

Bosnjak, Michael and Tracy L. Tuten, "Prepaid and Promised Incentives in Web Surveys: An Experiment," *Social science computer review*, 2003, *21* (2), 208–217.

Butler, Brian, "Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures," *Information Systems Research*, 2001, *12* (4), 346–362.

Buunk, Bram and Thomas Mussweiler, "New Directions in Social Comparison Research," *European Journal of Social Psychology*, 467-475 2001, *31* (5).

Cason, Timothy and Vai-Lam Mui, "Social Influence in the Sequential Dictator Game," *Journal of Mathematical Psychology*, 1998, *42*, 248–265.

Charness, Gary and Matthew Rabin, "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, August 2002, *117* (3), 817–869.

Chen, Yan, "Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research," in C. Plott and V. Smith, eds., *The Handbook of Experimental Economics Results*, Amsterdam: Elsevier Press, forthcoming.

Cosley, Dan, Dan Frankowski, Loren Terveen, and John Riedl, "SuggestBot: using intelligent task routing to help people find work in wikipedia," in "Proceedings of the 12th international conference on Intelligent user interfaces" 2007, pp. 32 – 41.

__ , __ , Sara Kiesler, Loren Terveen, and John Riedl, "SuggestBot: using intelligent task routing to help people find work in wikipedia," in "Proceedings of the SIGCHI conference on Human factors in computing systems" 2005, pp. 11 – 20.

__ , Pamela Ludford, and Loren Terveen, "Studying the Effect of Similarity in Online Task-Focused Interactions," *Proceedings of the 2003 International ACM SIGGROUP Conference on Supporting Group Work*, 2003, pp. 321–329.

Cox, James C., Daniel Friedman, and Steven Gjerstad, "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior*, April 2007, *59* (1), 17–45.

Dholakia, Utpal M., Richard P. Bagozzi, and Lisa Klein Pearo, "A social influence model of consumer participation in network- and small-group-based virtual communities," *International Journal of Research in Marketing*, September 2004, *21* (3), 241–263.

Duesenberry, James S., *Income, Saving and the Theory of Consumer Behavior*, Cambridge, MA: Harvard University Press, 1949.

Duffy, John and Nick Feltovich, "Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study," *International Journal of Game Theory*, February 1999, *28* (1), 131–152.

__ and Tatiana Kornienko, "Does Competition Affect Giving?," 2007. Working Paper, University of Pittsburgh.

Eckel, Catherine C. and Rick K. Wilson, "Social Learning in Coordination Games: Does Status Matter?," 2006. Working Paper, University of Texas at Dallas.

Falk, Armin and Urs Fischbacher, "A Theory of Reciprocity," *Games and Economic Behavior*, February 2006, *54* (2), 293–315.

Fehr, Ernst and Klaus M. Schmidt, "The Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, August 1999, *114* (3), 817–868.

Festinger, Leon, "A Theory of Social Comparison," *Human Relations*, 1954, *7*, 117–40.

Frank, Robert H., "The Demand for Unobservable and Other Nonpositional Goods," *American Economic Review*, March 1985, *75* (1), 101–116.

Frey, Bruno S. and Stephan Meier, "Social Comparisons and Pro-social Behavior - Testing 'Conditional Cooperation' in a Field Experiment," *American Economic Review*, December

2004, *94*, 1717–1722.

Goldberg, Lewis R., "The Structure of Phenotypic Personality Traits," *American Psychologist*, 1993, *48*, 26–34.

Groves, Theodore and John O. Ledyard, "Incentive Compatibility since 1972," in Theodore Groves, Roy Radner, and Stanley Reiter, eds., *Information, Incentives and Economic Mechanisms: Essays in Honor of Leonid Hurwicz*, Minneapolis: University of Minnesota Press, 1987.

Harper, Maxwell, Xin Li, Yan Chen, and Joseph Konstan, "An Economic Model of User Rating in an Online Recommender System," in Liliana Ardissono, Paul Brna, and Antonija Mitrovic, eds., *User Modeling 2005: 10th International Conference*, Vol. 3538 Springer Berlin/Heidelberg July 2005.

Harrison, Glenn W. and John A. List, "Field Experiments," *Journal of Economic Literature*, December 2004, *42* (4), 1009–1055.

Hopkins, Ed and Tatiana Kornienko, "Running to Keep in the Same Place: Consumer Choice as a Game of Status," *American Economic Review*, 2004, *94* (4), 1085–1107.

Hughes, Daniel, Geoff Coulson, and James Walkerdine, "Free riding on Gnutella revisited: the bell tolls?," *IEEE Distributed Systems Online*, June 2005, *6* (6).

Jones, Stephen R. G., *The Economics of Conformism*, Oxford & New York: Basil Blackwell, 1984.

Karau, Steven J. and Kipling D. Williams, "Social loafing: A meta-analytic review and theoretical integration," *Journal of Personality and Social Psychology*, 1993, *65*, 681–706.

Knez, Marc J. and Colin F. Camerer, "Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments," *Games and Economic Behavior*, July 1995, *10* (1), 65–94.

Krupka, Erin and Roberto Weber, "The Focusing and Observational Effects of Norms," 2005. Working Paper, Carnegie-Mellon University.

Levine, David K., "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, July 1998, *1* (3), 593–622.

Ludford, Pamela J., Dan Cosley, Dan Frankowski, and Loren Terveen, "Think different: increasing online community participation using uniqueness and group dissimilarity," in "Proceedings of the SIGCHI conference on Human factors in computing systems" 2004, pp. 631 – 638.

Pollak, Robert A., "Interdependent Preferences," *American Economic Review*, June 1976, *66* (3), 309–320.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, December 1993, *83* (5), 1281–1302.

Robson, Arthur J., "Status, the Distribution of Wealth, Private and Social Attitudes to Risk," *Econometrica*, July 1992, *60* (4), 837–857.

Rosnow, Ralph L., Robert Rosenthal, Roberta Marmer McConochie, and Robert L. Arms, "Volunteer Effects on Experimental Outcomes," *Educational and Psychological Measurement*, 1969, *29*, 825–846.

Samuelson, Larry, "Information-Based Relative Consumption Effects," *Econometrica*, 2004, *72* (1), 93–118.

Shang, Jen and Rachel Croson, "Field Experiments in Charitable Contributions: The Impact of Social Influence on the Voluntary Provision of Public Goods," 2005. Wharton School Working Paper, University of Pennsylvania.

Suls, Jerry, Rene Martin, and Ladd Wheeler, "Social Comparison: Why, with Whom, and with What Effect?," *Current Directions in Psychological Science*, 2002, *11*, 159–163.

Wash, Rick and Emilee Rader, "Public Bookmarks and Private Benefits: An Analysis of Incentives in Social Computing," in "ASIS&T Annual Meeting" 2007.

Table 1: MovieLens membership cohorts across treatments

Treatment	Membership Cohort	Total # users (active users)	Months in ML Mean	Std dev	min	max
Rating Info	New	45 (27)	3.1	1.1	0.2	5.5
	Mid	45 (35)	14.3	8.0	5.5	31.2
	Old	44 (37)	56.5	11.5	32.1	69.1
Net Benefit	New	44 (31)	3.2	1.3	0.2	5.5
	Mid	43 (27)	11.8	4.7	5.5	20.9
	Old	43 (32)	54.3	24.7	23.0	113.8
Control	New	55 (32)	2.9	1.2	0.9	5.5
	Mid	39 (25)	14.1	5.4	5.7	26.2
	Old	40 (31)	55.7	17.5	28.2	113.8

Table 2: Changes in ratings in Rating Info and Control: All (Active) Users

Treatment	$\Delta x_{E,i} = x_{E,i}^{t+1} - x_{E,i}^{t}$	New	Mid	Old	Overall
Rating Info	Below median: $\Delta x_{R,l}$	24.1 (51.7)	27.3 (58.4)	15.1 (20.6)	22.2 (39.9)
	Median Group: $\Delta x_{R,m}$	8.3 (20.8)	12.7 (14.7)	4.8 (5.6)	8.7 (12.4)
	Above median: $\Delta x_{R,h}$	-108.3 (-116.0)	15.4 (15.4)	-0.1 (-0.1)	-31.0 (-32.4)
Control	Below Median $\Delta x_{C,l}$	-3.3 (-6.0)	-1.8 (-4.6)	5.2 (8.5)	-0.3 (-0.7)
	Median Group $\Delta x_{C,m}$	-32.9 (-53.8)	7.3 (10.6)	3.6 (4.3)	-9.9 (-13.9)
	Above Median $\Delta x_{C,h}$	-22.3 (-38.5)	6.2 (7.4)	0.4 (0.5)	-7.5 (-10.2)

*

Table 3: Rating Info: Hypotheses and Wilcoxon Rank Sum Tests

All Users	Hypotheses	New	Mid	Old	Overall
	H_0: $x_l^{t+1} = x_m^{t+1}$	z = 1.29	z = -1.42	z = 0.11	z = 0.03
Below vs. Median	H_1: $x_l^{t+1} > x_m^{t+1}$	p = 0.10	p = 0.92	p = 0.46	p = 0.49
	H_0: $x_h^{t+1} = x_m^{t+1}$	z = 2.16	z = 0.35	z = 0.75	z = 1.96
Above vs. Median	H_1: $x_h^{t+1} < x_m^{t+1}$	p = 0.98	p = 0.64	p = 0.77	p = 0.97
	H_0: $x_m^{t+1} = x_c^{t+1}$	z= -0.92	z= 1.51	z= 0.65	z= 0.83
Median vs. Control	H_1: $x_m^{t+1} \neq x_c^{t+1}$	p= 0.36	p= 0.13	p= 0.52	p= 0.41
Active Users	**Hypotheses**	**New**	**Mid**	**Old**	**Overall**
	H_0: $x_l^{t+1} = x_m^{t+1}$	z = 2.29	z = 0.91	z = 0.77	z = 2.11
Below vs. Median	H_1: $x_l^{t+1} > x_m^{t+1}$	p = 0.01	p = 0.18	p = 0.22	p = 0.02
	H_0: $x_h^{t+1} = x_m^{t+1}$	z = -0.12	z = -0.16	z = 0.41	z = 0.19
Above vs. Median	H_1: $x_h^{t+1} < x_m^{t+1}$	p = 0.45	p = 0.44	p = 0.64	p = 0.58
	H_0: $x_m^{t+1} = x_c^{t+1}$	z= 0.20	z= 0.31	z= 0.33	z= 0.52
Median vs. Control	H_1: $x_m^{t+1} \neq x_c^{t+1}$	p= 0.84	p= 0.76	p= 0.74	p= 0.60

Table 4: Rating Info: Conformity of Two Types of Users

Panel A	Below	Median	Above	t-tests (1-sided)	
Type	x_l^{t+1}	x_m^{t+1}	x_h^{t+1}	$x_l^{t+1} > x_m^{t+1}$	$x_h^{t+1} < x_m^{t+1}$
Non-conforming	36.3	15.9	13.8	0.13	0.43
Conforming	59.7	18.5	20.6	0.03	0.63
Panel B	**Pre-experiment**		**Post-experiment**	**Paired sample t-tests (1-sided)**	
Type	$\lvert X_i^t - X_m^t \rvert$		$\lvert X_i^{t+1} - X_m^{t+1} \rvert$	$\lvert X_i^t - X_m^t \rvert \geq \lvert X_i^{t+1} - X_m^{t+1} \rvert$	
Non-conforming	317.9		313.6	0.20	
Conforming	401.8		394.4	0.04	

Table 5: Changes in Popular Movie Ratings in Net Benefit and Control: All (Active) Users

$\Delta x_i^p = x_i^{p,t+1} - x_i^{p,t}$	New	Mid	Old	Overall
Below average: Δx_l^p	0.9 (1.1)	-0.3 (-1)	13.1 (26.1)	4.5 (8.3)
Average: Δx_a^p	2.9 (5.7)	17.9 (22.8)	16.1 (20.5)	12.3 (17.8)
Above average: Δx_h^p	-28.9 (-36.2)	8.6 (10.8)	7.1 (7.6)	-4.4 (-5.2)
Control: Δx_C^p	-21.1 (-36.3)	3.8 (6.0)	1.6 (2.1)	-7.1 (-10.8)

Table 6: Contributions and Altruism for Above-Average Users

Activities	Altruism	All	Active	Hypotheses	P-values: All	Active
	Low	16.3	21.0	Low $>$ Middle	0.240	0.180
Rate Pop	Middle	11.0	13.1	Low $>$ High	0.290	0.210
	High	8.3	8.3	Middle $>$ High	0.380	0.300
	Low	0.7	0.9	Low $<$ Middle	0.134	0.148
Rate Rare	Middle	2.0	2.4	Low $<$ High	0.010	0.024
	High	5.8	5.8	Middle $<$ High	0.037	0.066
	Low	1.0	1.3	Low $<$ Middle	0.386	0.416
Database	Middle	1.3	1.6	Low $<$ High	0.167	0.225
	High	3.8	3.8	Middle $<$ High	0.120	0.165

Table 7: Contributions and Demographics in Treatments

D.V.	Ratings (x_i^{t+1})		Pop Ratings		Rare Ratings		Database Entries	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Users	All	Active	All	Active	All	Active	All	Active
Pre-rating	-0.001	-0.013	0.003	-0.077	1.653	0.632		
	(0.038)	(0.033)	(0.094)	(0.086)	(1.246)	(1.175)		
Above	2.351	-9.592	-12.610	-18.800**	-4.296	-8.419	10.310**	10.100**
	(11.150)	(10.460)	(7.956)	(7.440)	(6.939)	(6.858)	(3.985)	(3.986)
Below	26.740**	35.650***	-13.770	-11.390	-17.700**	-17.430**	2.872	3.433
	(11.870)	(11.580)	(8.643)	(8.633)	(8.270)	(8.546)	(4.834)	(4.916)
ML Age	-0.030	-0.095**	0.007	-0.023	0.017	-0.007	0.022	0.020
	(0.046)	(0.044)	(0.029)	(0.027)	(0.025)	(0.024)	(0.014)	(0.014)
Male	1.461	-3.569	7.069	0.119	-5.779	-8.871	6.734	7.087
	(11.930)	(11.30)	(7.400)	(7.476)	(6.658)	(6.900)	(4.384)	(4.280)
Age	2.128	9.579*	2.457	5.905	5.595*	8.561**	-1.553	-1.840
	(5.053)	(5.095)	(3.799)	(3.899)	(3.285)	(3.498)	(2.306)	(2.256)
Education	-2.300	-1.503	-0.353	0.181	-0.424	-0.410	-2.841***	-2.831***
	(1.937)	(1.929)	(1.300)	(1.185)	(1.170)	(1.119)	(1.051)	(1.049)
Student	-22.990	-25.330	-1.114	0.591	-0.637	0.317	3.313	2.448
	(16.550)	(15.710)	(10.600)	(9.961)	(8.941)	(8.784)	(3.939)	(3.932)
CompMath	-9.737	-8.977	-10.410	-14.510*	-11.600	-12.400*	3.306	2.766
	(11.560)	(10.690)	(8.182)	(7.640)	(7.591)	(7.392)	(3.509)	(3.503)
EdTLib	-23.940	-26.190*	-17.650	-16.050	-20.870*	-16.560	6.110	6.245
	(16.640)	(15.700)	(10.900)	(10.720)	(11.350)	(11.210)	(6.020)	(5.921)
Constant	46.600	30.530	11.170	10.410	-13.500	-12.200	25.410*	27.480*
	(34.880)	(32.940)	(25.140)	(23.960)	(22.110)	(21.660)	(13.830)	(13.800)
Obs.	100	83	95	74	95	74	95	74
L. L.	-409.3	-385.5	-334.1	-309.5	-206.6	-196.7	-56.4	-54.5
Pseudo R^2	0.014	0.029	0.013	0.024	0.039	0.043	0.214	0.200

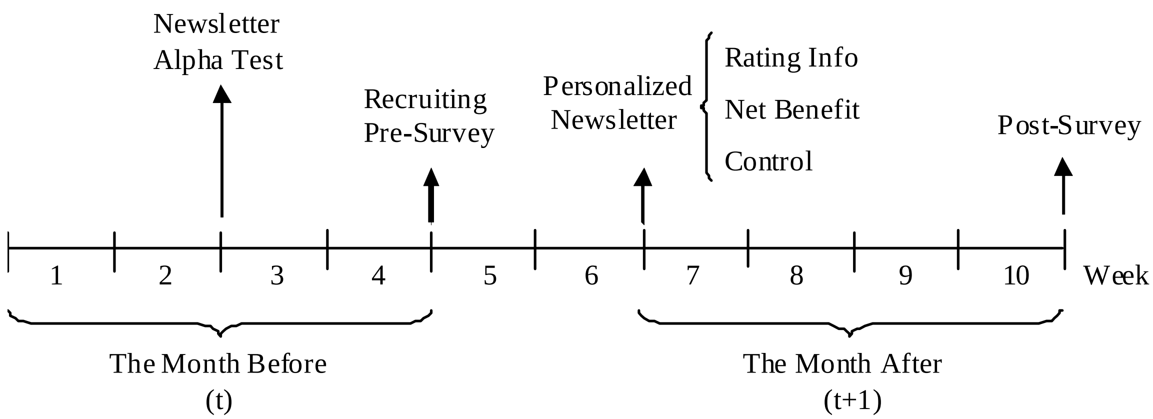


Figure 1: Experiment time line

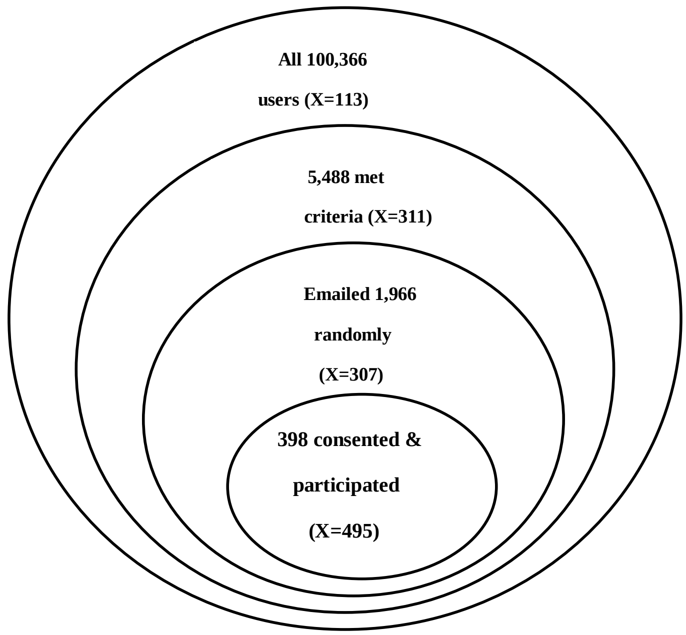


Figure 2: Sample and Population Comparisons

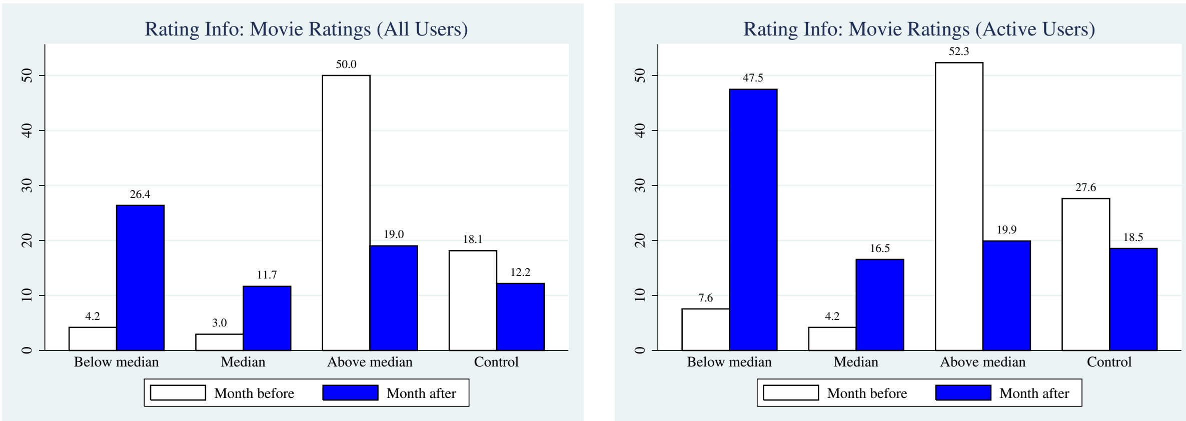


Figure 3: Rating Info treatment and control: Per user rating activities

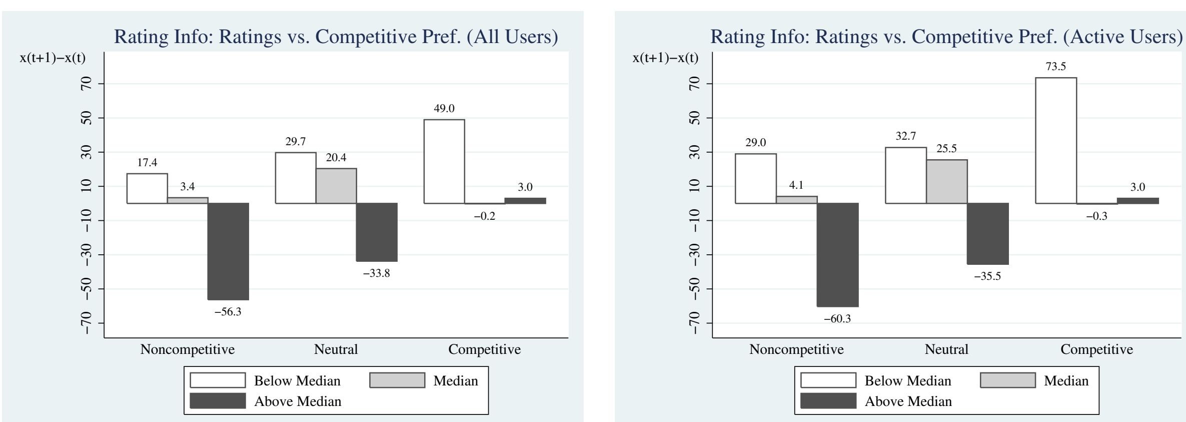


Figure 4: Change in ratings and self-reported competitiveness

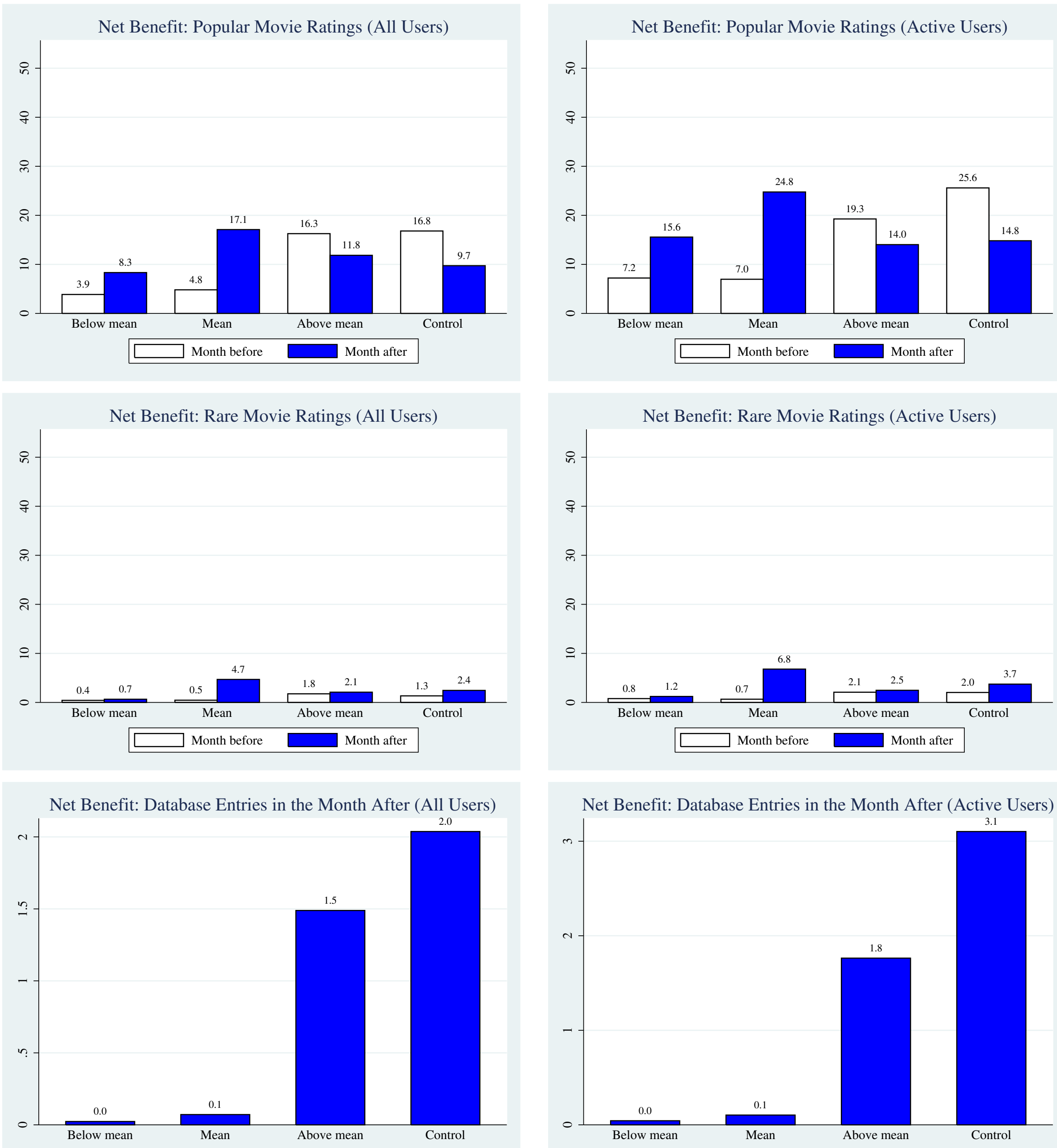


Figure 5: Net Benefit Treatment: Per User Activities