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

Yan Chen
Univ. of Michigan

Maxwell Harper
Univ. of Minnesota

Joseph Konstan
Univ. of Minnesota

Sherry Li
UT-Dallas

(www.communitylab.org)

November 13, 2008

Outline

- Online communities and under-contribution problem
  - MovieLens (www.movielens.org)
  - Recommender systems
- Social comparison theory
- Experimental design
- Theory
- Analyses and main results
- Discussions

Online Communities

- Online communities:
  groups of people meet to
  - Share information: e.g. cancer support groups
  - Produce info goods: e.g. open source
  - Play games: e.g. ESP games
  - Carry out business: e.g. Xerox service engineers

- Opportunities to create new social capital

Free-rider Problem

- Online communities:
  groups of people meet to
  - Share information: e.g. cancer support groups
  - Produce info goods: e.g. open source
  - Play games: e.g. ESP games
  - Carry out business: e.g. Xerox service engineers
- Nonparticipation and under-contribution (Butler 2001)
  - 50% of the social, hobby and working mailing lists had no traffic over a 122 day period

Example: Napster/Kazaa

- 66% of users share absolutely nothing (2000*)
- 1% of users share 50% of the content
- 25% of users share 98%
- By 2005, the free-rider problem is worse
  - 85% of users share nothing[†]

* Adar and Huberman, "Free Riding on Gnutella", First Monday, 2000
† Hughes et al., "Free Riding in Gnutella Revisited: the bell tolls", 2005

5

Under-contribution: Solutions

- Incentive-compatible mechanisms for public goods provision
  - Tax-subsidy schemes
- Online communities
  - Rarely use monetary transfers
  - Voluntary participation
  - Voluntary contribution
  - Lots of information about users
- Social information as non-pecuniary mechanism:
 - Social comparison theory

Social Comparison Theory

- **Festinger (1954):** people evaluate themselves by comparison with other people
- **Social comparisons affect behavior** (Suls, Martin and Wheeler 2002)
  - Information for the right behavior
  - Ambiguous situations
- **Conformity theory**
  - Akerlof (80)
  - Jones (84)
  - Bernheim (94)
- **Inequality aversion**
  - Fehr and Schmidt (99)
  - Bolton and Ockenfels (00)

Social Information in the Lab and Field

- Lab experiments
  - Cason and Mui (1998): sequential dictator game
  - Duffy and Feltovic (1999): learning
  - Bohnet and Zeckhouser (2004): ultimatum games
  - Krupka and Weber (2005): binary dictator game
- Field experiments
  - Frey and Meier (2004): mail fundraising
  - Shang and Croson (2005): on-air fund drive
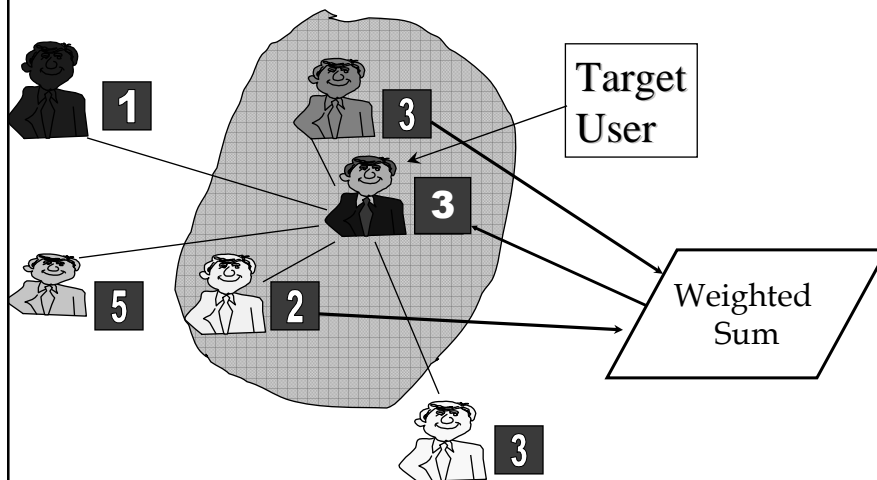
Designer's Challenge

- Peripheral participants
  - become active contributors
- Core participants/power users
  - Sustain and improve contribution
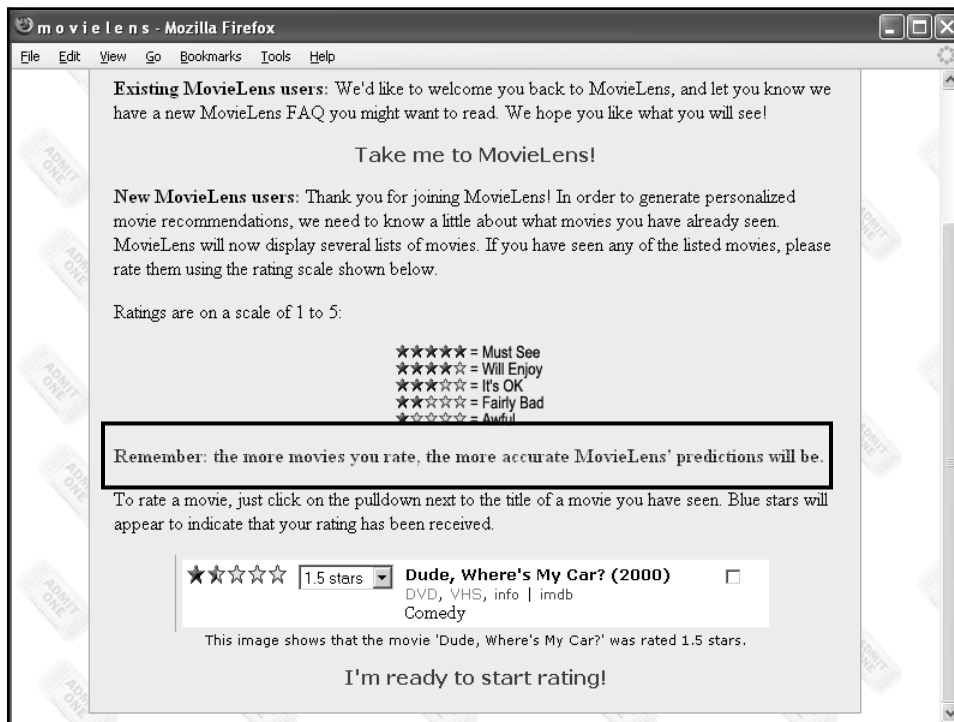- Personalized social information
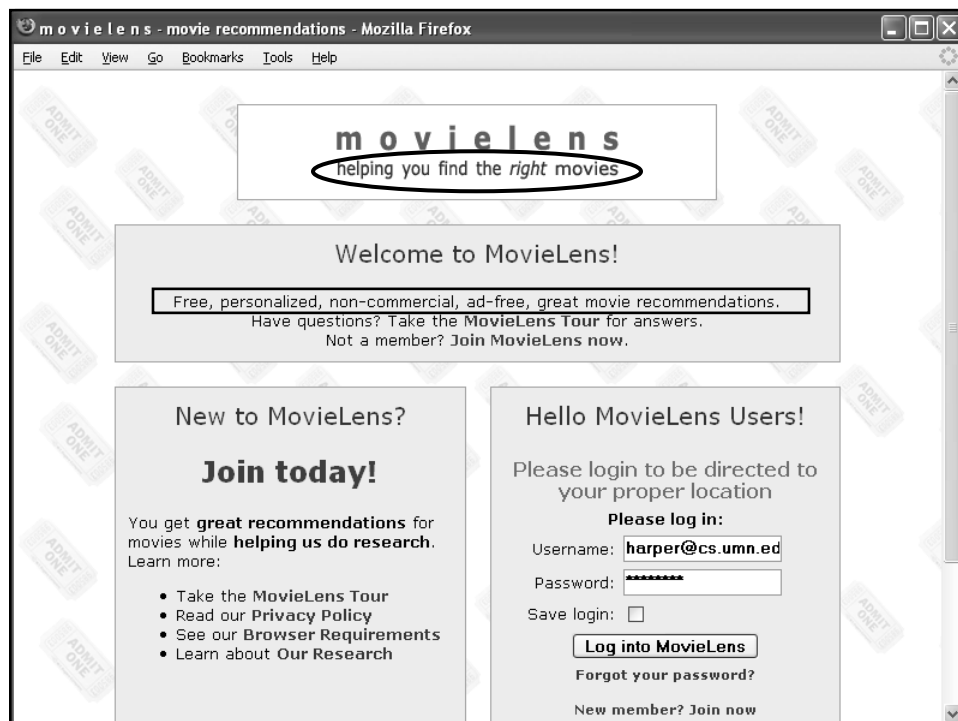
movielens.org

- Active and successful online communities
  - 100,000 users, 15,000 active within the past year
  - 13 million ratings of 9,043 movies
- Main activities
  - Rate movies
  - Receive recommendations
- Collaborative filtering technology
- 22% of movies have few than 40 ratings
- Software can't make accurate predictions

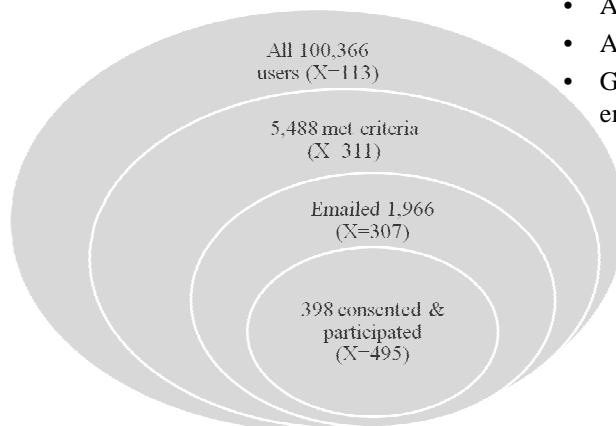
K-Nearest Neighbor Collaborative Filtering
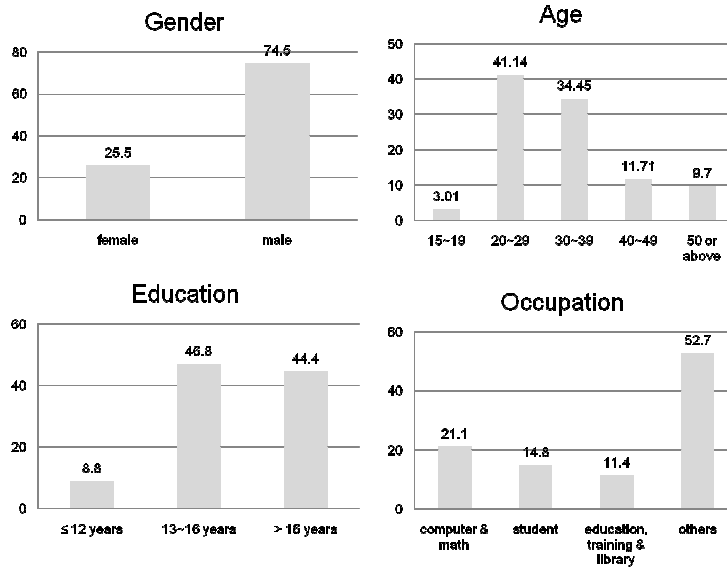

Source: John Riedl

Screenshot of MovieLens in Mozilla Firefox:

Existing MovieLens users: We'd like to welcome you back to MovieLens, and let you know we have a new MovieLens FAQ you might want to read. We hope you like what you will see!

Take me to MovieLens!

New MovieLens users: Thank you for joining MovieLens! In order to generate personalized movie recommendations, we need to know a little about what movies you have already seen. MovieLens will now display several lists of movies. If you have seen any of the listed movies, please rate them using the rating scale shown below.

Ratings are on a scale of 1 to 5:

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

Remember: the more movies you rate, the more accurate MovieLens' predictions will be.

To rate a movie, just click on the pulldown next to the title of a movie you have seen. Blue stars will appear to indicate that your rating has been received.

★★☆☆☆ [1.5 stars ▼] **Dude, Where's My Car? (2000)** ☐
DVD, VHS, info | imdb
Comedy

This image shows that the movie 'Dude, Where's My Car?' was rated 1.5 stars.

I'm ready to start rating!

Movielens Users and Our Sample

All 100,366 users (X̄=113)

5,488 met criteria (X̄ 311)

Emailed 1,966 (X=307)

398 consented & participated (X̄=495)

Criteria:
- Active in the past year
- At least 30 ratings
- Given us permission to send emails

398 Participants

Gender

female 25.5
male 74.5

Age

15~19 3.01
20~29 41.14
30~39 34.45
40~49 11.71
50 or above 9.7

Education

≤12 years 8.8
13~16 years 46.8
>16 years 44.4

Occupation

computer & math 21.1
student 14.8
education, training & library 11.4
others 52.7

Experimental Design

- Stage 1: Pre-experiment survey (398/1966)
  - Time to search for and rate ten movies
  - Willingness to pay for a list of top-ten movies
  - Number of ratings: perceived position
  - Net benefit: perceived position in distribution
- Stage 2: Experimental Newsletter
  - RatingInfo treatment: 134 users
  - NetBenefit treatment: 130
  - Control: 134

Pre- survey

Personalized newsletter

RatingInfo
NetBenefit
Control

Week 1   2   3   4   5   6
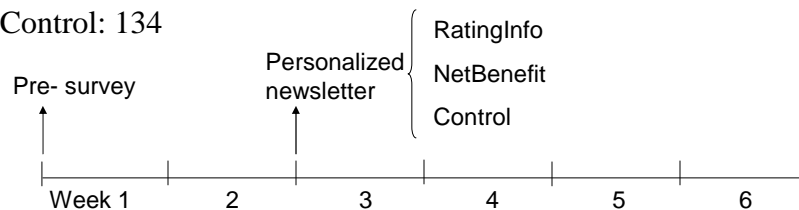
Experimental Design

- Stage 3: post-experiment survey
  - ML related questions
  - General social survey
  - Personality
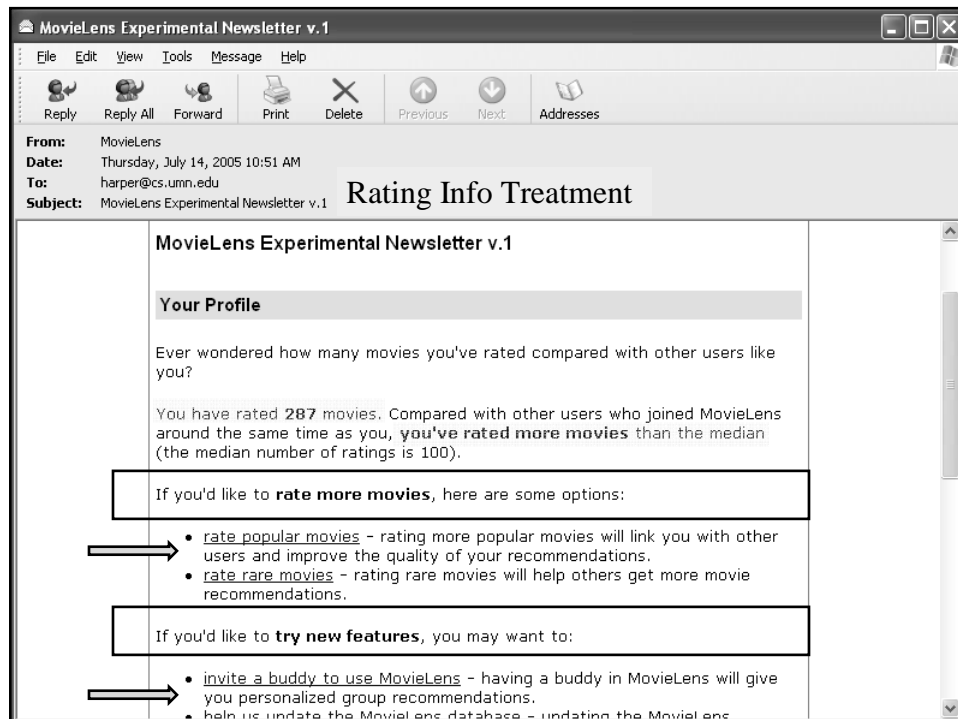 - Demographics
- Survey response rate: 78%

Pre- survey		Personalized newsletter	RatingInfo NetBenefit Control		Post- survey
Week 1	2	3	4	5	6
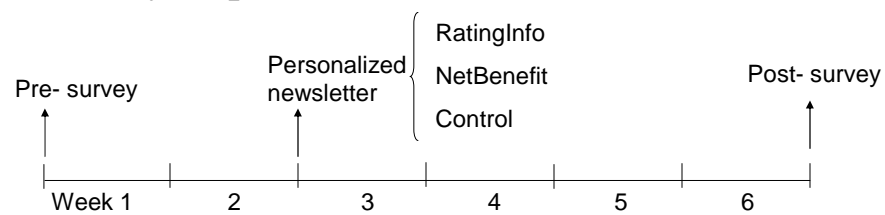
Stage 2: Social Information in Newsletter

- RatingInfo treatment
  - median number of ratings by similar users
  - Groups: Below-median, median, above-median

- NetBenefit treatment
  - average benefit score of similar users
  - Groups: Below-average, average, above-average

- Control
  - Percentage of the movies that you've rated are comedies…

Stage 2: Options of Activities

- Same five shortcuts for each condition
  - Rate popular movies: increase own benefit, easy
  - Rate rare movies: costly, but help other users
  - Update database: costly, but help other users
  - Invite a buddy: increase own benefit, easy
  - Just visit the Movielens homepage

MovieLens Experimental Newsletter v.1

File   Edit   View   Tools   Message   Help

Reply   Reply All   Forward   Print   Delete   Previous   Next   Addresses

From: MovieLens
Date: Thursday, July 14, 2005 10:51 AM
To: harper@cs.umn.edu
Subject: MovieLens Experimental Newsletter v.1

Rating Info Treatment

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 more movies** than the median (the median number of ratings is 100).

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

movielens
helping you find the *right* movies
Now with 8,715 movies, 96,940 users, and 11,931,422 ratings!

MovieLens Exp Net Benefit Treatment

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
helping you find the *right* movies
Now with 8,715 movies, 96,940 users, and 11,931,422 ratings!

MovieLens Experimental New Control

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.

m o v i e l e n s
helping you find the *right* movies

Welcome Max (Log Out)
You've rated 321 movies.
You're the 21st visitor in the past hour.

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

Home | Forums | Manage Buddies | Your Account | Help

Shortcuts | Search

You've searched for **all titles**.
Found **8435** movies, sorted by **Number of Ratings**
Genres: **All** | Exclude Genres: **None**
Dates: **All** | Domain: **All** | Format: **All** | Languages: **All**

Activity page: rate popular movies

Search by Genre

All Genres ▾   All Dates ▾

Domain: All movies ▾

☐ Use selected buddies!

Search Genre!

Advanced Search

Select Buddies

☐ Cindy
☐ Yahoo

What are buddies?

(hide) Predictions for you ↘	Your Ratings	Movie Information	Wish List
★★★↵	Not seen ▾	**Dances with Wolves (1990)** DVD info|edit|details|imdb Adventure, Drama, Western	☐
★★★★	Not seen ▾	**Apollo 13 (1995)** DVD info|edit|details|imdb Drama	☐
★★★↵	Not seen ▾	**Independence Day (ID4) (1996)** DVD info|edit|details|imdb Action, Adventure, Sci-Fi, War	☐
★★★★	Not seen ▾	**Aladdin (1992)** DVD info|edit|details|imdb Animation, Children, Comedy, Musical	☐
★★★★	Not seen ▾	**Fargo (1996)** DVD info|edit|details|imdb Crime, Drama, Thriller	☐
★★★	Not seen ▾	**Ace Ventura: Pet Detective (1994)** info|edit|details|imdb Comedy	☐

---

m o v i e l e n s
helping you find the *right* movies

Welcome Max (Log Out)
You've rated 321 movies.
You're the 21st visitor in the past hour.

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

Home | Forums | Manage Buddies | Your Account | Help

Shortcuts | Search

You've searched for **all titles**.
Found **8435** movies, sorted by **Date Added**
Genres: **All** | Exclude Genres: **None**
Dates: **All** | Domain: **All** | Format: **All** | Languages:

Activity page: rate rare movies

Search by Genre

All Genres ▾   All Dates ▾

Domain: All movies ▾

☐ Use selected buddies!

Search Genre!

Advanced Search

Select Buddies

☐ Cindy
☐ Yahoo

What are buddies?

(hide) Predictions for you ↘	Your Ratings	Movie Information	Wish List
???	Not seen ▾	**Acacia (2003)** DVD VHS info|edit|details|imdb Horror - Korean *Added 2005-07-31*	☐
★★★↵	Not seen ▾	**Race with the Devil (1975)** DVD info|edit|details|imdb Action, Horror, Mystery *Added 2005-07-31*	☐
???	Not seen ▾	**Three Stooges Meet Hercules, The (1962)** DVD info|edit|details|imdb Comedy, Fantasy *Added 2005-07-31*	☐
???	Not seen ▾	**Summer Magic (1963)** DVD info|edit|details|imdb Children, Comedy, Musical *Added 2005-07-31*	☐
???	Not seen ▾	**Best of Everything, The (1959)** DVD info|edit|details|imdb Drama, Romance *Added 2005-07-31*	☐

Activity page: update database

Check Suggested Movie: Last Hurrah for Chivalry

Here's the information we received:

Title: Last Hurrah for Chivalry
Starring:
Directed by:
Language(s): English
Genre(s): (no genres listed)

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)

Creating Peer Groups Based on ML Age

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	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	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

A Theoretical Framework:
A Neoclassical Model Harper, Li, Chen, Konstan (2005)

- User's *neoclassical* benefit function:

$$\pi_i(X_i, \sum_{j\neq i} X_j) = \underbrace{\gamma_i Q_i(X_i, \sum_{j\neq i} X_j)}_{\substack{\text{benefit from} \\ \text{recommendations}}} + \underbrace{f_i(X_i)}_{\substack{\text{rating} \\ \text{fun}}} + \underbrace{v_i(d_i)}_{\substack{\text{enjoyment} \\ \text{from} \\ \text{updating} \\ \text{database}}} + \underbrace{h_i}_{\substack{\text{fun from} \\ \text{browsing}}} - \underbrace{c_i(X_i)}_{\substack{\text{cost} \\ \text{of} \\ \text{rating} \\ \text{movies}}} - \underbrace{c_i^d(d_i)}_{\substack{\text{cost of} \\ \text{updating} \\ \text{database}}}$$

where Q quality of recommendations, γ_i marginal benefit from movie recommendation, X number of ratings, d database entries,

$f' > 0, f'' \leq 0, v' > 0, v'' \leq 0, c' > 0, c'' \geq 0.$

- Ratings: private and public good

Solution and Model Estimation

- Solution: inefficient amount of rating

$$\ln x_i^* = \frac{1}{1-\beta_i}[\ln \beta_i + \ln \gamma_i - \ln(c_i - f_i)]$$

- Model estimation

$\ln x_i = a_0 + a_1\gamma_i + a_2\beta_i + a_3 f_i + a_4 c_i + \overline{\Lambda}\overline{Z} + \varepsilon_i$

$\ln x_i = 4.351 + 0.028\gamma_i + 0.971\beta_i + 0.353 f_i - 0.042 c_i + \overline{\Lambda}\overline{Z} + \varepsilon_i$

  – Explains 34% of variance in rating behavior

Extension to a 2-period model

- t: the month before pre-survey
- $t+1$: the month after newsletter
- $X_{i:}$ user i's life time rating
- x_i: user i's monthly rating
- d_i: user i's number of database entries
- Without social information: neoclassical model
- With social information:
  - Conformity
  - Difference aversion

Rating Info Treatment

- Extending to 2-period model and incorporating conformity

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

- g_i : disutility from deviating from social norm
- Proposition 1
  - Below-median users rate more movies than median users: $x_l^{t+1} \geq x_m^{t+1}$
  - Above-median users rate less than median users $x_h^{t+1} \leq x_m^{t+1}$
  - Conformity to median $|X_i^{t+1} - X_m^{t+1}| \leq |X_i^t - X_m^t|$

Net Benefit Treatment

- Incorporating distribution preferences

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

$$\text{average score users}: \quad u_a^{t+1} = \pi_a^{t+1}$$

$$\text{low score users}: \quad u_l^{t+1} = \pi_l^{t+1} - g_l(\pi_a^{t+1} - \pi_l^{t+1})$$

$$\text{high score users}: \quad u_h^{t+1} = \pi_h^{t+1} - g_h(\pi_h^{t+1} - \pi_a^{t+1})$$

- g_l envy; g_h charity concern
- **Proposition 2**
  - Average/low score users will rate more popular movies
  - High score users will rate rare movies and update database if g_h is sufficiently large

---

Results (RatingInfo)
Below-median users increase ratings by 530%.
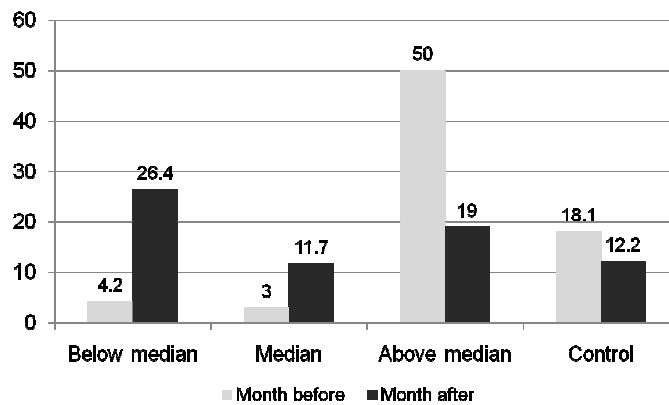
- Below-median users rated more than median (p=0.02);
- Above-median users rated more than median (p=0.03);
- Distance from median is smaller in the month after (p=0.02).

Conformity and Competition

- Below-median users: +530%
- Above-median users: - 60%
- Asymmetry in behavior change may be due to competitiveness
- Survey measure on competitiveness
  - "It's achievement, rather than popularity with others, that gets you ahead nowadays."

Conformity and Competition
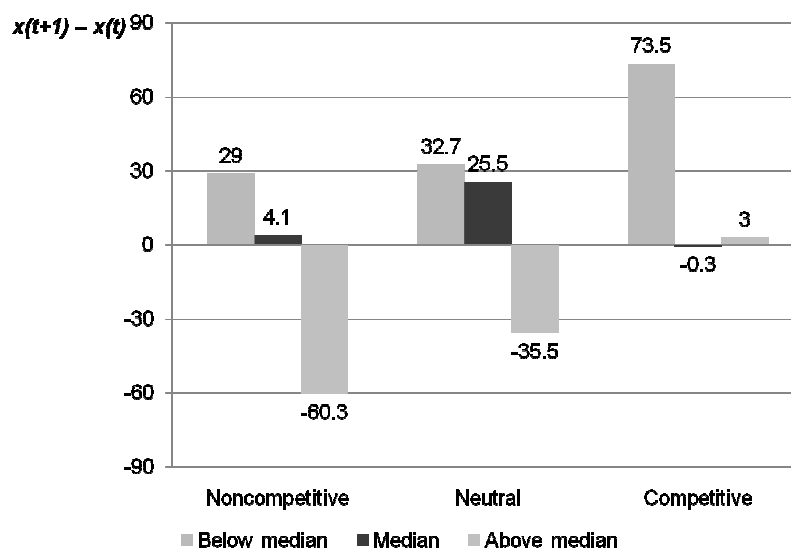
Net Benefit Treatment: Popular Ratings

- Above < Average ; Below < Average (p=0.03)
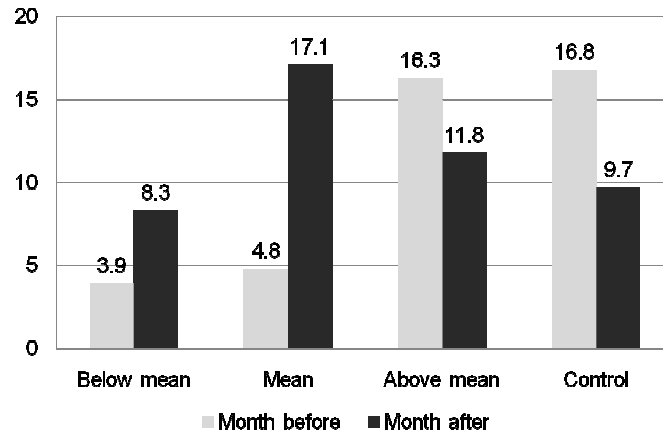
Net Benefit Treatment: Rare Ratings

- Above > Below (p=0.01); Average > Below
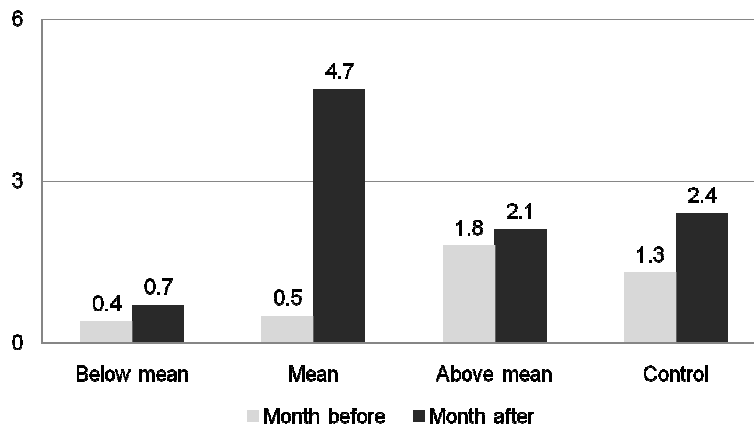
Net Benefit Treatment: Database Entries

- 94% of database entries come from above-mean users
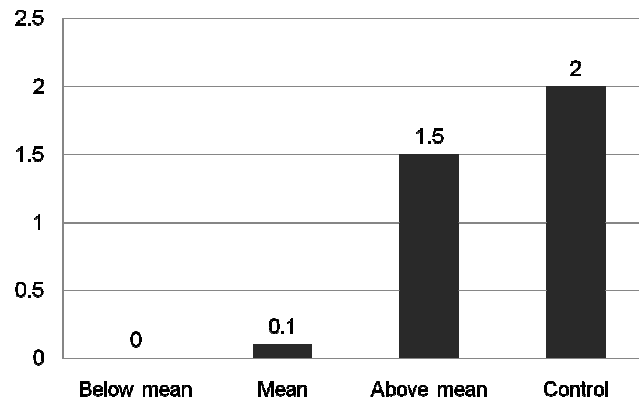
Measures on Altruism

- Prop 2: High-benefit score users will rate rare movies and update database if *charity concern* is sufficiently large
- Survey: "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."

High Net Benefit Users:

More altruistic users rated less popular movies
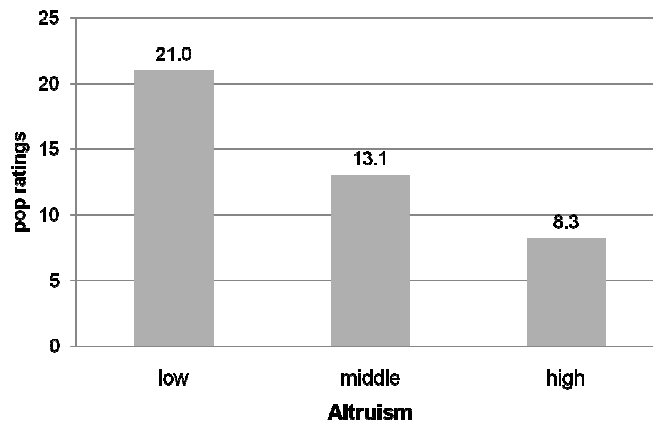
High Net Benefit Users:

More altruistic users rated more rare movies
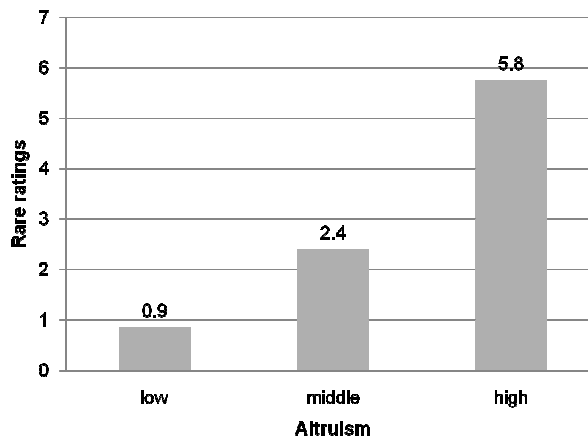
High Net Benefit Users:

More altruistic users had more database entries
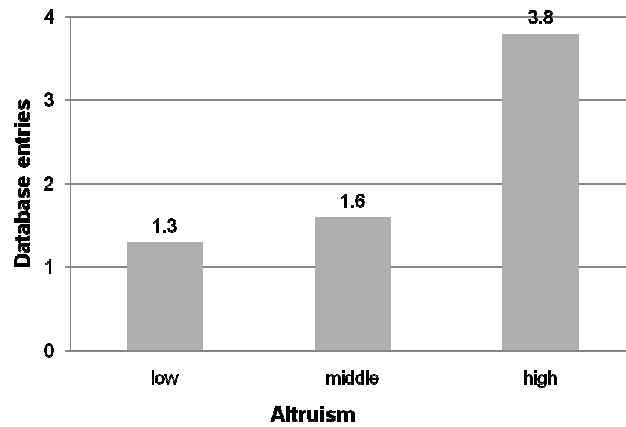
Are the effects due to anchoring?

	Information provided	Corr (x^{t+1}, info); (p values)
Rating Info	You have rated xxx movies	-0.091 (0.297)
	Median number of ratings	-0.058 (0.503)
Net Ben	Your net benefit score	0.070 (0.428)
	Average net benefit score	0.144 (0.103)
Control	% ratings that were comedies	-0.135 (0.119)
	Average rating in this genre	0.139 (0.110)

P.S. Regression to the mean? No. (Difference in difference analysis)

Red Queen Effect

- The Red Queen said, "… it takes all the running you can do, to keep in the same place."

         – Lewis Carroll's *Through the Looking-Glass*

- Rating Info:

  relative rankings of total movie ratings remain the same (Spearman > 0.9)

- Net Benefit:

  relative rankings of net benefit scores remain the same (Spearman >= 0.8)

Summary

- Social comparison significantly influence behavior
- Rating Information
  - Below median: increase # of ratings by 530%
  - Above median: decrease # of ratings by 62%
- Conformity vs. competitive preferences
- Net Benefit
  - Above : rate more rare; update more database entries
  - Average: rate more pop
  - Effect of altruism
- Design Implications: personalization

Design Implications: Personalization

- Personalized social information
  - Below
    - median rating information
    - More ratings
  - Power users
    - Average net benefit
    - High-cost activities that help others

Future Work

- Other forms of social information
  - Leaderboard: ESP game
- Other reward
  - Promotion: slashdot
  - Barnstar: wikipedia
- Work-oriented online communities
  - SourceForge

Knowledge Market Design:
A Field Experiment at Google Answers

Yan Chen, Yong-Mi Kim, Teck Ho

February 13, 2009

Outline

- Background:
  Internet Question-and-Answer Services
- Design problem:
  contract design for knowledge markets
- Experimental design
- Results
- Discussions

Internet Q&A Services

- Commercial services: use money
  - Google Answers (2002-2006)
  - Uclue (2005-present)
  - BitWine
- Community-based services
  - AnswerPoint (from AskJeeves.com, ask.com)
  - Yahoo! Answers
  - AnswerBag

Internet Q&A Services

Table 1: Internet Q&A Services

Name	No. questions	Who answers	Price & Tip	Reputation system
Google Answers	53,087	Researchers selected by Google	$2 to $200	1 to 5 stars
Yahoo! Answers	10 million+	All registered users	No	Points, levels
Answerbag	365,000+	All registered users	No	Points, levels
Internet Public Library	50,000+	Librarians and LIS students	No	No

Notes:

1. Google Answers number of questions includes only those that can still be accessed through their archive. Some estimates have placed number of questions around 150,000.
2. According to Yahoo!'s blog, Yahoo! Answers had their 10 millionth answer posted on May 7, 2006.

Google™

Answers

Ask a question. Set your price. Get your answer.

More than 500 carefully screened Researchers are ready to answer your question for as little as $2.50 -- usually within 24 hours. Your satisfaction is completely guaranteed. Find out what others are saying.

Log in or
Create a Google Account

Step 1 - Enter your Question. Tips for great results.

Ask Question

Search Google Answers for [____] [all questions ▼] Google Search

Browse previously asked questions

Arts and Entertainment   Reference, Education and News

Business and Money   Relationships and Society

Computers   Science

Family and Home   Sports and Recreation

Health   Miscellaneous

Recently answered questions

Fire Departments participating in N...

Legal obligation

MY KEYBORED IS STUK IN UPPPERCASE

Gambling on the internet in Hong Ko...

view all questions

Google Home - Answers Help & Tips - Answers FAQ - Terms of Service - Privacy Policy
©2006 Google

Done

Google Answers View Question

Log in | Google Answers Home

Ask a Question

Q: Retirement financing statistic in Canada (Answered ★★★★★, 0 Comments)

Sponsored Links:

Government CA Income Tax	Neighborhood Information	401K Retirement Plan	Disability Benefit Canada
Articles and Information about Government Of Canada Income Tax	Get House Value, Income Level, Average Age, Crime Level, More	Helpful Links for 401K Retirement Plan	Disability Benefit Canada Options! For Disability Benefit Start Here.
Tunu.com	www.Intelius.com	www.ToSeekA.com	Disability.Benefit.AlltheBrands.com

Question

Subject: **Retirement financing statistic in Canada**
Category: Business and Money
Asked by: babushkaya-ga
List Price: $50.00

Posted: 01 Apr 2005 08:42 PST
Expires: 01 May 2005 09:42 PDT
Question ID: 503655

Greetings,

I am looking for a statistic on percentage (or absolute number) of people having RRSP accounts in general and self-managed RRSP accounts as a sub-set study. I am also curious on how those numbers change throughout years and if there is a tendency.

The second part of the question would be statistic on how many people have DB (define benefits) plans with their Canadian employers and how this number is changing throughout years.

Thank you.

Sponsored Links:

Demographic Statistics	Unemployed in canada	Self Employed 401k	Defined Benefit Plans
Create Custom Reports w/ Maps From The 2000 Census Data. Free Trial!	Loans For The Unemployed: When Job Loss Threatens Stability	Learn more about the Self Employed 401k retirement plan.	The Hartwood Group. Benefit Plan Specialists.
www.DemographicsNow.com	eresourcesforsuccess.com/Fin46.html	www.bcmadvisors.com	www.HartwoodGroup.com

Answer

Subject: Re: Retirement financing statistic in Canada
Answered By: wonko-ga on 04 Apr 2005 09:33 PDT
Rated: ★★★★★

"Four in five RPP members belong to a defined benefit plan"

"Pension Plans in Canada" by Pensions and Wealth Surveys Section, Statistics Canada (January 1, 2003) http://www.statcan.ca/english/research/13F0026MIE/13F0026MIE2004001.pdf

"Nearly two-thirds of Canadian taxfilers aged 20 to 59 contributed to a registered retirement savings plan (RRSP) at least once during the 1990s."

"Overall, the number of people contributing to RRSPs increased by more than 25% between 1993 and 2001."

"Study: using RRSP's before retirement" Statistics Canada (December 15, 2004) http://www.statcan.ca/Daily/English/041215/d041215d.htm

"Canadians increased their contributions to a Registered Retirement Savings Plan (RRSP) in the 2003 tax year for the first time in three years. The number of contributors declined slightly."

Features of Google Answers

- User pricing: $2 - $200
  - GA provide pricing guideline
- Researchers select questions
- Reputation system
  - Consumer rating of answers: 1-5 stars
  - Total # of questions answered
  - Total # of refunds
  - Unique identification of researchers
- Social preference
  - Reciprocity: tips
  - Altruism: free comments, some community aspects

Literature: Field Data

- Economics
  - Edelman (2004)
    - More experienced answerer: higher rating
    - Positive corr between rating and answer length
  - Regner (2005)
    - Frequent askers are more likely to tip: reputation
    - 18% one-time user also tips: social preference
    - Answerers put more effort if asker tipped frequently before

Literature: Field Data

- **Information Science**
  - Rafaeli, Raban and Ravid (2005, 2007)
    - Data: GA 06/2002-10/2004
    - participation, i.e., # of questions answered
    - Average price of questions $19.37
    - Average dollar value of an answer: $20.20
  - Adamic, Zhang, Bakshy and Ackerman (2008): YA
    - 3 categories of forums
    - Answer length: significant predictor of best answers

Literature: Field Data

- **Information Science**
  - Yang, Adamic and Ackerman (2008)
    - Taskcn
    - Requester posts task and price
    - Sealed bid all-pay auction: winner gets money
    - Price uncorrelated with number of submissions
  - Nam, Ackerman and Adamic (2008)
    - Knowledge-iN
    - Interviews
    - Motivations for participation: altruism, personal learning and personal interests

Literature: Field Experiment

- Raban and Harper (2008)
  - Several sites
  - Quality
    - GA > free sites
    - YA > other free sites
    - Higher price leads to higher quality

Research Questions

- Knowledge market design
  - What's the effect of money?
    - Fixed price vs. bonus contracts
    - Effects of pricing on quality of answers
    - Traffic and amount of frivolous questions
  - What's the effect of a reputation system?
  - How social preference might change the optimal contract
    - Reciprocity
    - Altruism: community-based systems

Gift Exchange:
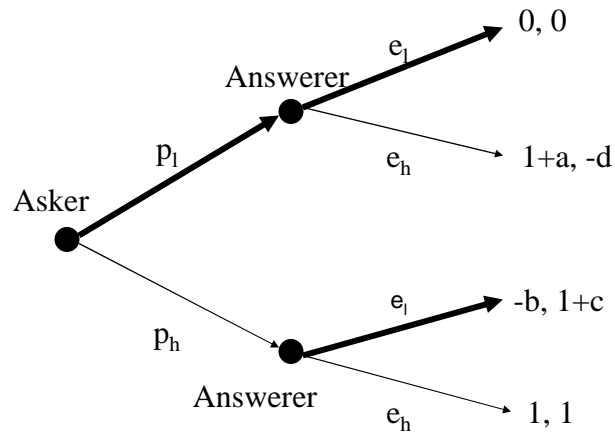
Sequential Prisoners' Dilemma: Dominant strategy equilibrium, SPNE

Gift exchange: reciprocal types
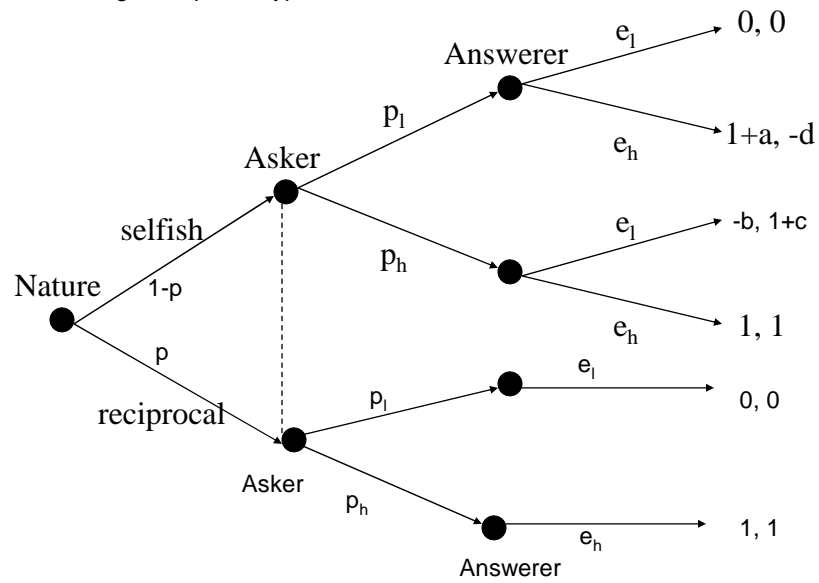
When p is sufficiently high, p_h and e_h can be supported in equilibrium.

Experimental Design

- Question selection
  - Real reference questions from IPL: 50,000
  - From 2003 on: time used on record
  - Select c.a. 1-hour questions, with verification by RA
  - Open ended: effort and quality can vary

- Problem with question selection
  - Our censorship might have removed the variations

- Compare with Harper et al (2008)

Experimental Design

- Four treatments (based on 10,000 GA questions)
  - $20 fixed price (baseline)
  - $30 fixed price
  - $20 + $10 tip conditional on good answer
  - $20 + $10 tip unconditional

- Sufficient variations in price?

- Timing: sent 4 per day, one from each category

- One-shot: new user for each question
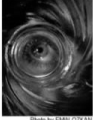
ipl The Internet Public Library

Search the IPL
Now you can "Grok" the IPL by selecting "Grok It"!
Then hit "Map View" for a visualization of your results.
Advanced Search
Search ⦿ Traditional (fast) ○ Grok It (clustered)

Accessibility Info
Subject Collections
Arts & Humanities
Business
Computers
Education
Entertainment
Health
Law & Government
Regional
Science & Tech
Social Science
Ready Reference
Almanacs
Calendars
Dictionaries
More Ready Reference...
Reading Room
Books
Magazines
Newspapers
KidSpace
TeenSpace
Special Collections
Blogs
Literary Criticism
Science Fair
More Special Collections...
Searching Tools
FARQs
Pathfinders
Search this Site
Web Searching
About the IPL
IPL Audio/Video Tours
Teach with the IPL
Contact Us
Ask a Question

Spotlight

Bottled Up?

Photo by EMIN OZKAN

The IPL **teenspace** Poetry Wiki is a great place for you to express yourself. Post your poems & poetic inspirations or respond to others' work.
Release your inner poet...

See All of the IPL's Spotlight Features...

Inside

An Exciting New Era for the IPL!

The IPL Consortium

The IPL was founded by a class at the University of Michigan's School of Information, and Michigan SI students almost exclusively generated its content and managed the Ask a Question reference service. Soon, 14 other schools will be actively taking part in the development and maintenance of the IPL.

Read about how the IPL Consortium came to be...

See who will be taking part...

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

The Internet Public Library is a public service organization and a learning/teaching environment at the University of Michigan School of Information.

sponsored by
intel ❖ Sun

Examples: IPL Questions

- Our school theme this month is APPRECIATION. Can you recommend any stories or books that emphasize that theme for Middle School students (grade 6-8)? I also need for elementary students, but Middle School is a priority.

- I have heard that supposedly when after you receive a massage you are often told to drink lots of water (or fluids, I would guess) because receiving a massage releases toxins and you need to flush them out of your system. Frankly, this sounds like hippie stuff. Is there any actual hard scientific evidence to support or refute this claim? If your answer is good, I'll add a $10 tip.

10K Questions from GA

Table 2: Summary Statistics from 10K Downloaded Q&A from Google Answers

Price Range	% answered	% adding tip	mean price	median price	tip/price	OBS
[$0, $5]	38.2	13.2	3.3	2.6	1.21	4570
($5, $10]	36.6	19.6	7.2	7.0	0.52	2077
($10, $25]	36.0	17.0	17.8	20.0	0.42	2078
($25, $100]	39.0	19.5	46.0	50.0	0.29	1380
($100, $200]	45.8	19.6	180.2	200.0	0.20	212
[$20, $40]	34.9	18.2	24.4	23.6	0.35	1871
[$0, $200]	37.7	16.2	18.4	10.0	0.71	10317

Hypotheses

- Price effects
  - Higher price leads to longer answers (effort)
  - Higher price leads to better answers (quality)

- Tip effects
  - Promise of a tip leads to longer answers (effort)
  - Promise of a tip leads to better answers (quality)

- Reputation
  - Researchers with higher past average rating will provide higher quality answers

Data Analysis

- Outcome measures
  - Length of answers: word count
  - Quality of answers: rater data
- Control
  - Difficulty of each question
- Raters
  - Inter-rater reliability test
  - Semi-professionals:

    SI graduate students finished SI 647

Rating Procedures

- 16 raters from SI 647
  (Information Resources and Services)
- Two sets of 100 GA question & answer pairs
  - Set A: 75 IPL + 25 GA
  - Set B: 100 GA from 10K questions
- Each q&a rated by 8 independent raters
- Training sessions + 5 rating sessions
- Randomized order of q&a

http://www-personal.si.umich.edu/~kimym/questions.php

Q: Retirement financing statistic in Canada

Question

Subject: **Retirement financing statistic in Canada**   Question ID: 503655

Greetings,

I am looking for a statistic on percentage (or absolute number) of people having RRSP accounts in general and self-managed RRSP accounts as a sub-set study. I am also curious on how those numbers change throughout years and if there is a tendency.

The second part of the question would be statistic on how many people have DB (define benefits) plans with their Canadian employers and how this number is changing throughout years.

Thank you.

Answer

Subject: **Re: Retirement financing statistic in Canada**

"Four in five RPP members belong to a defined benefit plan"

"Pension Plans in Canada" by Pensions and Wealth Surveys Section, Statistics Canada (January 1, 2003)
http://www.statcan.ca/english/research/13F0026MIE/13F0026MIE2004001.pdf

"Nearly two-thirds of Canadian taxfilers aged 20 to 59 contributed to a registered retirement savings plan (RRSP) at least once during the 1990s."

"Overall, the number of people contributing to RRSPs increased by more than 25% between 1993 and 2001."

"Study: using RRSP's before retirement" Statistics Canada (December 15, 2004)
http://www.statcan.ca/Daily/English/041215/d041215d.htm

"Canadians increased their contributions to a Registered Retirement Savings Plan (RRSP) in the 2003 tax year for the first time in three years. The number of contributors declined slightly."

"Registered retirement savings plan contributions" Statistics Canada (November 2, 2004)
http://www.statcan.ca/Daily/English/041102/d041102b.htm

"In total, Canadians had accumulated an estimated $1.15 trillion in the three main retirement programs - RPPs, RRSPs, and the Canada and Quebec Pension Plans (C/QPP) - by the end of 2001. This was almost double the level of $593.6 billion in 1990, when measured in constant (inflation-adjusted) dollars. Of the total assets in 2001, 69% were in RPPs, 25% in RRSPs (individual and group), and about 6% in the C/QPP."

"Study: Workers' knowledge of retirement plans" Statistics Canada (January 23, 2004)
http://www.statcan.ca/Daily/English/040123/d040123b.htm

The following resource provides a number of statistics on pension plans and RRSPs that may be of interest to you:

"Canada's retirement income programs" Statistics Canada (November 17, 2003)
http://www.statcan.ca/Daily/English/031117/d031117a.htm

"Table 111-0039 summarizes Registered Retirement Savings Plan (RRSP) contributions, by contributor characteristics at the national,

Please rate the difficulty of the **question**
(1=very easy ... 5=very difficult):
○ 1   ○ 2   ○ 3   ○ 4   ○ 5

Please rate the **answer** for the following factors
(1=strongly disagree ... 5=strongly agree):

2a. The question that was asked is answered.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2b. The answer is thorough, addressing all question parts.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2c. The sources cited are credible and authoritative.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2d. The links provided are to relevant pages.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2e. Information in the cited sources is summarized.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2f. Only information pertinent to the question is presented.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

2g. The answer is well-organized and written clearly, avoiding jargon and/or inappropriate language.
○ 1   ○ 2   ○ 3   ○ 4   ○ 5   ○ NA

3. Please rate the overall quality of the answer (1=very low quality ... 5=very high quality):
○ 1   ○ 2   ○ 3   ○ 4   ○ 5

Submit Query

Rating answer quality

1. Please rate the difficulty of the question

2. Please rate the answer for the following factors:
   A. The question that was asked is answered.
   B. The answer is thorough, addressing all question parts.
   C. The sources cited are credible and authoritative.
   D. The links provided are to relevant web sites or pages.
   E. Information in the cited sources is summarized.
   F. Only information pertinent to the question is presented.
   G. The answer is well-organized and written clearly, avoiding jargon and/or inappropriate language.

3. Please rate the overall quality of the answer

Interrater Reliability

- Interrater reliability provides an indication of the extent to which variance in the ratings is attributable to differences among objects rated
  - i.e. relation of one rated object to other rated object is same across judges → high interrater reliability
- Intraclass correlation coefficient
 - Multi-rater generalization of Cohen's Kappa

Interrater Reliability

Group	Difficulty (Q1)	Overall Quality (Q3)	Summed (Q2 A-G)
A (IPL)	0.71	0.77	0.78
A (Non-IPL)	0.86	0.77	0.73
A (all)	0.77	0.77	0.77
B	0.89	0.72	0.72

Interrater reliability assessed with ICC[3,8]

Analysis of Multi-Item Scale

- Q2 parts A-G - summated multi-item rating scale measuring "quality"
  - Cronbach's alpha = 0.84
  - Items reflect single underlying construct
- Pearson correlation of summated scale (Q2 A-G) and overall rating (Q3)
  - Group A: 0.75 to 0.92
  - Group B: 0.74 to 0.95
  - Q2 A-G and Q3 measuring something very similar
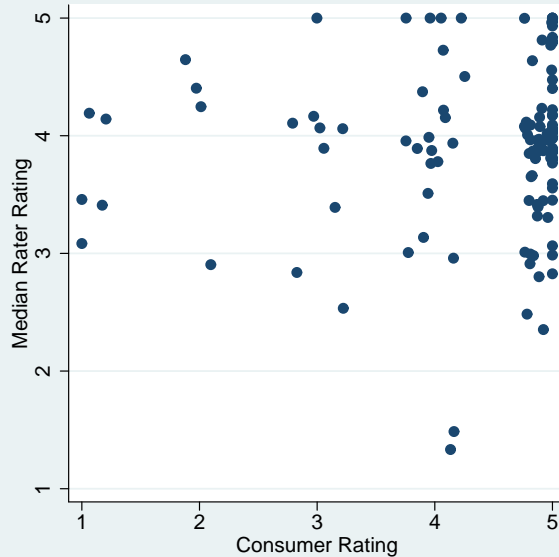
Rater rating = 0.847 Consumer rating (p<0.01), based on 125 GA questions.

Distribution of Average Quality (Raters)

Result: Determinants of Answer Length

- Answer length as a proxy for effort:
  - corr(reported time, wordcount)= 0.635
  - p=0.015, n=14

- Regression results:
  - Higher price leads to significantly longer answers
  - Researchers with higher past reputation gave significantly longer answers

Tobit Regressions

Table 4: Determinants of Answer Length (Effort)

	Dependent Variable: Word Count		
	(1) IPL	(2) GA	(3) All
Price	7.472	13.097	12.575
	(23.035)	(2.545)***	(2.128)***
Tip	16.862	25.519	27.115
	(21.164)	(19.357)	(13.796)*
Reputation	1,368.709	1,073.285	1,143.625
	(434.286)***	(613.795)*	(413.810)***
Experience	-0.244	-0.136	-0.168
	(0.128)*	(0.126)	(0.095)*
Constant	-5,083.371	-4,259.175	-4,448.396
	(2,011.381)**	(2,687.905)	(1,801.455)**
Observations	75	125	200

Result: Determinants of Answer Quality

Ordered Probit

	Dependent Variable: Quality Rating		
	(1) IPL	(2) GA	(3) All
Price	-0.035	-0.000	-0.001
	(0.035)	(0.002)	(0.002)
Tip	-0.009	0.001	0.005
	(0.032)	(0.016)	(0.013)
Reputation	1.358	0.742	0.996
	(0.670)**	(0.501)	(0.396)**
Experience	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	75	125	200

Notes:

a. Ordered probit: standard errors are in parentheses.

b. Significant at: * 10-percent; ** 5-percent; *** 1-percent level.

Quality: Reputation Effects

- Price is not a significant predictor of quality
- Nor is an ex ante promised tip
- Researcher past reputation is the most significant predictor for quality
- Compare with Harper et al (2008)
  - Rater background: MSI vs English
  - Official answer vs. [answer+comments]

Summary

- Effort (word count):
  - Higher price leads to longer answers and more links
  - Higher reputation leads to longer answers
- Quality
  - Not price elastic
  - Higher reputation leads to better answers

Knowledge Market Design

- Money
  - Eliminates frivolous questions
  - Reduces volume
  - Higher prices leads to longer, but not better answers: pay enough or don't pay at all?
- Reputation
  - Greater effort and higher quality
- Point and level system: making contribution evaluable, but multi-dimensional tasks, might skew incentives

Thanks …