How Social Motivation Enhances Economic Activity and Incentives in the Google Answers Knowledge Sharing Market

Sheizaf Rafaeli, Daphne R. Raban, Gilad Ravid

Center for the Study of the Information Society (InfoSoc), University of Haifa,
Mount Carmel, Haifa, 31905, Israel
Department of Industrial Engineering and Management, Ben Gurion University of the Negev,
P.O.Box 653, Be'er Sheva 84105, Israel

Sheizaf@rafaeli.net; draban@gsb.haifa.ac.il; rgilad@bgu.ac.il

Abstract

What are the antecedents, inhibitors and catalysts to providing information and participating in mixed fee-based and free online contexts? We describe the behaviour of about 500 "Researchers" in the Google Answers online service and over 70,000 questions posed on this system over a 29 month period. Google Answers is a fee-based environment. Answers provided on it are "worth" over $20 (including tips) on average. However, labour economics of response to price and tip alone do not fully account for the online information provision market. Non-monetary incentives, such as feedback in the form of comments predict and explain some of the variance in participation. Descriptive and correlational findings presented here are based on many thousands of answers. We thus corroborate some of the hybrid theories of information provision behaviour presented to date mostly in laboratory settings. The participation of experts in Google answers is associated with a hybrid of material (economic) and social motivators.

Keywords: Information markets, social incentives, economic incentives, knowledge sharing

Introduction

Information is either a commodity, to be treated as a simple economic good, or a more abstract, amorphous term describing anything from documented facts to myths and speculations. How can markets form around information? Disentangling this conundrum should be of broad interest. The willingness to sell, buy or share information has been studied in laboratory settings (Rafaeli, Raban et al. 2003; Rafaeli and Ravid 2003; Borck, Frank et al. 2006; Raban and Rafaeli 2006). Clearly, on the web, sharing and trading information co-exist. Information, is transferred by market transactions, or shared voluntarily through free giveaways or contributions. Sometimes, the business model is mixed, and is often confusing. For instance, some news is sold when it is fresh and given away for free when outdated. On the other hand some daily newspapers publish free copy on the web, while charging a fee for archived news. The hybrid form of information transfer, combining fee and free, calls for research scrutiny. What is the effectiveness of the unusual, hybrid, market form? What would attenuate actors' information behaviour in such real markets?

Information is expensive to produce and cheap to reproduce (Bates 1989; Shapiro and Varian 1999). The cost of information can be either direct or indirect. The quest for the right pricing for information and participation is further complicated by the fact that information is an experience good, meaning that its value is revealed only after consumption (Shapiro and Varian 1999; Van Alstyne 1999).

Behavioural research revealed that the value of information is derived from perceptions of at least three central elements: cost, quality, and ownership (Toften and Olsen 2004; Raban and Rafaeli 2006). Manipulation of one or more of these elements can have dramatic effects on information trading and sharing markets. And where manipulation is possible, social and cultural concerns come into play.

The Google Answers information market is a promising testbed for an empirical examination of this issue. Google Answers was established in April 2002. Information and monetary transactions in this global arena provide a large public database for research. Unlike other markets, here the focus is entirely on information. Google Answers combines commercial and voluntary transfer of questions, comments and answers. Thus, it is a market for both information and knowledge. In the following we describe the system, and then we explain the social incentives that are proposed as partial explanation for the inclination to provide answers in this system.

Google Answers (http://answers.google.com) is a fee-based information market where experts sell their expertise to askers for a price quoted by the askers (between $2 - $200 per question). Free sharing of information in the form of comments takes place alongside the information trades. Google Answers encompasses “Researchers” (or “Experts”) who provide responses (or “Answers”) to questions that have an associated Price, and post-answer Tip, and Ratings. Each answer may be preceded and/or followed by a discussion and comments. Persons who post questions ("askers"), must furnish a credit card and commit to a payment. Designated and certified responders (experts/researchers) are members of a pre-approved set, though they, like the askers, remain anonymous and hidden behind a pseudonym. Askers, experts and the general public may all participate in the discussion free-for-all.

This paper analyzes and reports on the relationships between participation, reward and feedback mechanisms on the Google Answers (GA) system. Data on the commercial transactions as well as the communication process include, beyond prices and sales, information about discussion in the form of textual comments, “star” ratings on a 1-5 scale provided by recipients, gratuity (tips paid to the responders), and the like. Statistical analyses of these data shed light on some of the interesting theoretical and ideological questions surrounding the value of information.

“Answers” are defined as responses given in response to specific questions. All published answers represent the payment of a fee. "Comments", on the other hand, constitute free advice, opinions and discussion that appear on the site, but are not associated with a monetary transaction, and do not necessarily involve either a paying customer or an accredit "Responder". The rules of the site call for “askers” (buyers) to provide questions accompanied by a pre-declared “Price”. Buyers commit to paying this price if and when their question is answered. The expert who provides the answer receives three quarters of the declared price and Google receives the rest.

The community in question here is carefully and formally circumscribed. While any owner of a valid credit card may post a question, only pre-approved “Experts” may provide a paid, sanctioned answer. Questions posted to Google Answers are publicly viewable on the Google Answers website. Any registered user can add their insights and share the benefit of the research. Answers are posted publicly as well. Askers and responders do not get to move into secluded corners, all transactions of money, information and feedback occur in plain sight. Users who provide comments are not paid for their posts, but they may add interesting perspectives to the data gathered by the Researcher. The identity and personal information of participants is not revealed at any time; All participants are

identified only by a self-selected Google Answers 'Nickname'. This fact, alone, makes for an interesting limit or flavour for the motivation to participate.

Two prior empirical projects focused on Google Answers (Edelman 2004; Regner 2005). In both cases, the theoretical orientation stressed the view of information as a commodity, and to varying degrees studied the behaviour of agents on Google Answers as an instance of labour economics. The guiding question driving previous research was an attempt to explain information economics in either labour or behavioural terms. This study is an examination of the economics of information from a social and communication perspective.

The theoretical question driving our investigation relates to the social motivations of participants in online forums in general, and fee-based, public information markets such as Google Answers in particular.

In earlier treatments of the same question, Edelman (2004) and Regner (2005) approached the question of incentives to participation as a labour economics problem. They found that experienced researchers received higher ratings, that answerers adjust their behaviour over time to better suit asker preferences, that the hourly pay for being active on the site as well as tips did positively predict the amount of effort invested (in other words, participation), and that experienced answerers were more specialized. However, in Edelman's data a counter-intuitive finding of interest in the context of incentives to participation was that more specialized answerers earned less per-hour. Edelman explains that when a researcher insists on staying within a particular substantive field he/she forgoes opportunities in other fields. This behaviour is cast as a lack of versatility and is therefore a negative characteristic on predicting earnings. Edelman also finds a labour economics (pay per time) perspective in differential compensation for times of day and days of week responses. For example, the “graveyard shift” is less popular, less rewarded, and less desirable.

Table 1 in the Results section provides descriptive statistics comparing the two earlier studies with the present research.

Our theoretical approach to this data set is to study the relations between economic, social and psychological incentives we have previously reviewed (Raban, Ravid et al. 2005; Rafaeli, Raban et al. 2005). We are interested in the value (incentive) added to participation by the social and communication arrangements. An initial inspection suggested a correlation between economic incentive (price) and amount of questions answered. Tips were only very weakly correlated, and the socially constructed ratings were not correlated at all. However, after the dataset was

pruned to contain only those question-and-answer pairs (and attendant “discussions”) in which at least one comment was provided, the correlations of socially-based incentives for participation rose, and became significant. In the following, we further investigate the relation between the discussion comments and the inclination to provide answers in order to establish the social contribution to site activity and, eventually, to economic activity.

Ling et al. (2005) review social psychological incentives to participation. They follow in the footsteps of Rafaeli and Larose (1993), Constant et al. (1994) and Kollock and Smith (1996) in expressing the group and communication based inputs that can be fed-back by the system in order to increase contribution, fidelity, commitment and sense of belonging. Following Ahituv (1989) and Rafaeli and Raban (2003) we search for a richer description and a quantitative measure of the confluence of the economic and behavioural to participation and the valuation of information.

The issue of demand for information or the willingness to pay for it has been addressed by behavioural as well as economic research. In economic terms information can be either a public or a private good. As a public good, few people pay for information but everyone enjoys it (Tragedy of the Commons). Exchanges in online forums are often cited as public goods. Private information goods require direct payment by each user. Information markets are unique in that the public and private information goods may co-exist adjacently and are therefore likely to affect each other's consumption patterns. The present research aims to investigate this relationship by looking at the association between free comments and for-fee answers.

The research hypothesis is, therefore, that the valuation of information, as revealed by market data, is not fully explained by labour economics or behavioural antecedents. The prices set by the market, and the motivation to act on these prices, are probably an outcome of more intricate dynamics, including influences that are based on the social milieu of information markets. This study is an attempt to unearth and document such influences. One departure point for clarifying this question is in the choice of unit of analysis. As the market's behaviour involves choices, it would probably be useful to analyze participants' behaviour, beyond inert item costs.

Method

Using a specially developed Perl web agent we gathered data that represents all the questions, answers and other content and transactions on the site. We parsed the text and inserted it into an SQL database for further analysis. The

web agent tool was designed to produce sequence URLs and fetch them rather than crawl the site. Using this method we were able to find unlinked pages, not just the information presented through menus . We produced a large database of questions, answers (for a fee), comments (for free) and all technical data available for these items including: nickname and ID for each asker and Expert, exact timing of questions, answer, and comment, price, tip, rating.

We collected all of Google Answers' site activity (questions, answers, comments etc.) since it's inception in April 2002 through December 7th, 2004. We removed all incomplete observations from the beginning and end of our sampling period and obtained a sample of 77,673 questions.

The resulting database of Google Answers activity was analyzed in two levels: the question/answer level, and the Expert level. In other words, some analyses were done for the entire set of questions or answers, and some analyses were done on summary data of the entire set of answered questions at the expert level.

Results

Table 1 provides descriptive statistics comparing the present study with earlier studies by Edelman (2004) and Regner (2005).

GA element	Current Study	Edelman (2004)	Regner (2005)
Period of Study	06/2002 – 10/2004	04/2002 – 11/2003	07/2003-01/2004
Duration	29 months	20 months	7 months
Number of questions asked	77,673	43,262	13,948
Number of answers provided	37,970	24,290	6,853
Number of questions with comments only	21,828	NA*	NA
Number of questions with comments	39,436	NA	NA
Number of comments sent	97,802	NA	NA
Rated answers	23,868	NA	NA
Tipped answers	7,503	NA	1,745
Number of experts	512	534	NA
Average dollar value of question	$19.37	NA	$19.23
Average dollar value of answer	$20.10	$18.91	$21.59
Average dollar value of unanswered question	$18.66	NA	NA
Average answer rating (on a 5 point scale)	4.60	4.33	4.70
Average answer tip value	$8.86	$8.77	$8.94
System price range	$2-200		
System tip range	$1-100		

* NA: Data not available in the article

Table 1. Descriptive statistics in Edelman (2004), in Regner (2005) and in the present study

The mean question price asked for was about 19.37 dollars. Neat, “round” figures were most popular, with over 12,000 of the questions priced at $5, over 13,000 priced at $10, and so forth. The standard deviation of question price was over 30. The mean tip (gratuity) was $8.86.

Of the 77,673 questions in our dataset, about one half, 37,970, were answered. Those answers were associated with conversation (interactivity) in the form of comments, 41,882 comments were contributed to the answered questions. The upshot of these preliminary findings reported previously (Rafaeli, Raban et al. 2005) was that, when interaction was present (answers with comments), the social parameters of rating and comments contributed incentives to the formation of participation, beyond the role of economic incentives. In this paper we report further, fine-tuned analysis. First, we isolated only those questions which had answers and comments. We then limited this subset further to include only those question-answer (Q/A) sets where at least one comment was submitted before the answer was produced. This Q/A subset contains 7,024 observations. Of those, 4,188 Q/A sets had comments only before the answer was given.

In order to evaluate the effect of comments on the experts' inclination to provide answers we summarized the Q/A sets on a per-expert basis. In the entire set of answers 512 active experts are available with answer counts ranging between 1 and 1,897. In the subset where comments were given before answers we identified 417 experts with answer counts between 1 and 901. Table 2 summarizes the main descriptives for the complete set of answers and the subset of interest. For this table and subsequent analyses we defined a new variable, the ratio between the number of comments before the answer was given and the answer count. We call this variable CBPA and use it as an indicator for the activity intensity of comments before answers. For example, if 40 comments were submitted before an expert submitted a total of 10 answers, this would mean the CBPA ratio is 4. The mean of ratios is reported in Table 2. The rationale for this ratio is to avoid the obvious correlation between number of comments and number of answers per expert.

	All Answers	**Comments before Answers**
Number of Experts	512	417
Number of Q/A sets	37,970	7,024
Expert mean answer count	74.16 (S.D. 181.70)	37.03 (S.D. 75.97)
Mean price	$16.90 (S.D. 16.14)	$16.58 (S.D. 17.79)
Mean tip	$8.25 (S.D. 7.16)	$9.32 (S.D. 9.70)
Mean rating	4.34 (S.D. 0.71)	4.40 (S.D. 0.79)
Mean ratio of comment-per-answer CBPA	1.46 (S.D. 1.57)	1.76 (S.D. 1.63)

Table 2: Descriptives Per Expert in the Google Answers information market

A t-test on the answer count per expert yielded a significant difference between more-than-average CBPA and less-than-average CBPA per expert ($t=5.9$; $p < 0.01$). This provides further support to the correlational observation, namely that the comments given before the answers elicit more answers from the experts. More comments elicit more answers.

Additional evidence for the social interpretation of the results comes from analyzing the full Q/A set to see whether the mere presence of comments before answers is associated with more answers, not at the expert level, but at the Q/A level. For this purpose we ran a chi-square test. The chi statistic was very high and statistically significant ($p < 0.01$).

The following matrix was tested using the Chi-Square test:

	No: Answer	Yes: Answer	
No: Comment-before-Answer	17,893	33,782	51,675
Yes: Comment-before-Answer	21,810	4,188	25,998
	39,703	37,970	77,673

Table 3: Data matrix for the Chi-Square test

A multiple regression analysis was conducted on the entire set of experts (N=512) to predict the experts' tendency to provide answers from the two variables that were present and known to the experts prior to answering, namely price (the price bid by the asker) and the ratio of CBPA per answer per expert. The overall R^2 was 0.05 and the adjusted R^2 was 0.046. The main contribution to the explained variance came from the CBPA variable where R=0.213 ($p<.01$) while the correlation with price was R=0.081 ($p<.05$).

Discussion

Google Answers provides a natural field research opportunity and setting to study the value of information and the structure of online markets of information. As Google Answers is a public database it provides invaluable access to complete details of information market transactions such as trading (buying and selling answers) and sharing knowledge and advice. While the frequency of requests for refunds following an unsatisfactory answer is not known, most of the dynamics of the market are laid bare. The longevity of the system suggests that refund requests do not occur too frequently and are not a defining characteristic or detrimental to operations.

The system descriptive statistics raise interesting questions.

Why are only about half of the questions answered for a fee? Unanswered questions are not much cheaper than answered questions ($18.66 vs. $20.10). Thus, the exclusive economic explanation, that Experts are price-seekers is not necessarily supported.

Why do people contribute so many free comments in a system that is at essence a market and does provide an economic scaffold? In fact, the number of comments (i.e. freely provided information) is more than double the number of answers.

What accounts for tips in an anonymous information market? Why do anonymous askers choose to tip experts identified only by a nickname? Tips are generous, given in about 20% of the answers, with a mean of $8.86 for about 20% of the answers. Tips amounted to 4.42% of the income generated by selling answers. For a face-to-face environment such as a restaurant or hotel lobby this rate of gratuity may sound modest. However, this level of tipping may be considered high in a voluntary, anonymous online system with no prior traditions or set norms and with no accumulated reason to provide tips. Both sellers and buyers have little or no mutual acquaintance, less expected future relationship, and no identification. The system has no memory that can be affected by gratuities. Recall that (according to one urban myth) "tip" stands for "To Insure Promptness". Promptness is one thing this system has even without tips. One explanation for online tipping is that people tend to transfer the traditional economic behaviour with which they are familiar from the real world and apply it in online contexts despite the greater freedoms afforded online. We have found this previously in relation to the subjective value of information in trading and sharing contexts (Rafaeli and Raban 2003; Raban and Rafaeli 2006).

Tipping emerges as a promising solution to one of the central problems in the value of information, namely, that information is an experience good. As such its value is subjective and changes by circumstance and person (Ahituv 1989; Shapiro and Varian 1999). The real value of information is uncovered only following consumption and use, often after a time gap. This is the point where tips in an asynchronous system can compensate for the difference between the initial price bid and the final value to the consumer. Our previous research demonstrated an Endowment Effect in relation to information consumption patterns (Raban and Rafaeli 2006). This effect suggests that price bids for buying information underestimate its value especially when compared with the value to the seller, who already knows more about that same information and, perhaps, can better evaluate it. The gap in valuations between buyers and sellers is reduced by tips. Regner (2005) investigated tips in the Google Answers system. This research showed that tips increase the subsequent tendency of experts to provide answers. In other words, the experts' information economic behaviour is not myopic. Experts learn to see beyond the initial price bids and expect a certain degree of tipping to compensate them further.

The analysis suggests that, beyond economics and pricing, social incentives play an important role in generating market activity. Social activity is seen by the amount of comments posted on Google Answers. A coldly pure and rational actor would not be expected to post a voluntary comment on a commercial web site which offers payment for activity. Why, then, is the number of comments so large and what is their effect? The simple reply in the words of George Mallory in 1924 when asked about climbing Mount Everest is: "Because it is there". People contribute information on many web sites so why not contribute in Google Answers as well. It's another site to show off or prove one's knowledge and search skills, and, possibly, to voice an opinion. However, our analysis suggests more than that. The chi square analysis in table 3 shows that the mere presence of comments before answers reduces the likelihood of an answer. The logic is simple: if sufficient help was provided by a comment, there is less need or room for an answer. Ethical experts will not post a paid answer where an informative comment was submitted.

Analysis of the data summarized in the expert level raises the observation that questions that generate much discussion in the form of comments are more likely to be answered by an expert. Experts seem to be drawn to questions that generate much interest, comments, activity and tend to answer those questions more often. This may fill a social need but may also serve an economic purpose of enhancing the expert's reputation by getting exposure to more eyeballs. Moreover, the correlation between the comments-before-answers with the tip of those answers is higher (0.223, $p<.01$) than the correlation between the total comments in answered questions with the tip given to those answers (-0.090, $p<.05$). At least two explanations are possible. First, the comments may enhance the overall perceived quality of all knowledge provided, as answer or comments, so the asker becomes more inclined to tip. A second explanation may be that the asker feels some social pressure by the presence of the comment contributors which leads him/her to provide a tip as a social norm. Social interaction, in the form of free sharing of knowledge in the form of comments, especially before answers are given, generates more economic activity (answering), and is also accompanied by enhanced tipping.

Another possible explanation for the enhancement of market activity by social activity comes from the theory of social facilitation. The presence of other people causes arousal and improves performance on simple or dominant, well-learned, tasks (Zajonc 1965). Previous research has shown that social facilitation plays an important role in economic behaviour such as in shopping (Sommer and Sommer 1989; Sommer, Wynes et al. 1992) and in online auctions (Rafaeli and Noy 2002). Google Answers is unique in that the presence of others is exhibited only by the

comments they post online. This is asynchronous presence which may suggest that we are witnessing asynchronous social facilitation. This notion invites further research, possibly using an experiment to determine causality.

The Google Answers information market thus emerges as a social space, perhaps a community. Despite the anonymity, this web site proves to be more than just an impersonal commercial service. This is a community of total strangers connected only by technology forming an amorphous web of weak ties. One can only explain the interactions and transactions of information by relying on a hybrid theory. Other price and informal means of motivation and reward are required to account for the transaction of information. The network provides incentives and ties of both an economic and a social nature. Hybrid ties combine to make for an ongoing social system that wraps the market.

These ties, that comprise social structure/network, await further research. The web which sometimes seems like a gigantic shopping mall ruled by large corporations offers new ways of combining economic and social gratifications, and this is, perhaps, one of its most compelling strengths. Markets for information can exist and sustain relatively large volumes of transactions. These transactions are ruled by more than crass supply and demand models. Our findings suggest a hybrid dynamic model that weaves exchange, labour and social motivations into a sustaining fabric that accounts for the flow of answers desired by those who pose the questions. Rather than falling into the trap of simplistic, uni-dimensional explanations, designers of online arenas should make use of the more complex, hybrid explanation.

References

Ahituv, N. (1989). Assessing the value of information: Problems and approaches. Proceedings of the Tenth International Conference on Information Systems, Boston, MA.

Bates, B. J. (1989). Information as an economic good: A reevaluation of theoretical approaches. Mediation, Information, and Communication. B. D. Ruben and L. A. Lievrouw. New Brunswick, NJ, Transaction Publishers. **3:** 379-394.

Borck, R., B. Frank, et al. (2006). "An Empirical Analysis of Voluntary Payments for Information Goods on the Internet." Information Economics and Policy **18**(2): 229-239.

Constant, D., S. Kiesler, et al. (1994). "What's mine is ours, or is it? A study of attitudes about information sharing." Information Systems Research **5**(4): 400-421.

Edelman, B. (2004). "Earnings and Ratings at Google Answers." Retrieved March, 2005, from http://cyber.law.harvard.edu/people/edelman/pubs/GoogleAnswers-011404.pdf.

Kollock, P. and M. Smith (1996). Managing the Virtual Commons: Cooperation and Conflict in Computer Communities. Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives. S. Herring. Amsterdam, John Benjamins.

Ling, K., G. Beenen, et al. (2005). "Using Social Psychology to Motivate Contributions to Online Communities." Journal of Computer-Mediated Communication **10**(4): article 10.

Raban, D. R. and S. Rafaeli (2006). "The effect of source nature and status on the subjective value of information." Journal of the American Society for Information Science and Technology **57**(3): 321-329.

Raban, D. R., G. Ravid, et al. (2005). Paying for Answers: An Empirical Report on the Google Answers Information Market. AoIR 6.0, Chicago, IL.

Rafaeli, S. and R. J. LaRose (1993). "Electronic bulletin boards and "public goods" explanations of collaborative mass media." Communication Research **20**(2): 277-297.

Rafaeli, S. and A. Noy (2002). "Online auctions, messaging, communication and social facilitation: a simulation and experimental evidence." European Journal of Information Systems **11**(3): 196-207.

Rafaeli, S. and D. R. Raban (2003). "Experimental investigation of the subjective value of information in trading." Journal of the Association for Information Systems **4**(5): 119-139.

Rafaeli, S., D. R. Raban, et al. (2005). Social Cognition Online. The social net: The social psychology of the internet. Y. Amichai-Hamburger. Oxford, England, Oxford University Press.

Rafaeli, S., D. R. Raban, et al. (2005). Social and Economic Incentives in Google Answers. ACM SIGGROUP, Sanibel Island, FL.

Rafaeli, S., D. R. Raban, et al. (2003). Online Simulations in Management Education about Information and Its Uses. Educating Managers with Tomorrow's Technologies. C. Wankel and R. DeFillippi. Greenwich, CT, Information Age Publishing Inc.

Rafaeli, S. and G. Ravid (2003). "Information sharing as enabler for the virtual team: an experimental approach to assessing the role of electronic mail in disintermediation." Information Systems Journal **13**: 191-206.

Regner, T. (2005). "Why Voluntary Contributions? Google Answers!" CMPO Working Paper Series(Working Paper No. 05/115).

Shapiro, C. and H. R. Varian (1999). Information Rules: A Strategic Guide to the Network Economy. Boston, Harvard Business School Press.

Sommer, R. and B. A. Sommer (1989). "Social facilitation effects in coffeehouses." Environment and Behavior **21**: 651–666.

Sommer, R., M. Wynes, et al. (1992). "Social facilitation

effects in shopping behavior." Environment and Behavior **24**: 285–297.

Toften, K. and S. O. Olsen (2004). "The relationships among quality, cost, value and use of export market information: An empirical study." Journal of International Marketing **12**(2): 104-131.

Van Alstyne, M. W. (1999). A proposal for valuing information and instrumental goods. International Conference on Information Systems. Charlotte, NC.

Zajonc, E. (1965). "Social facilitation." Science **149**: 269-274.