Three essays on the economics of information systems

by

Lian Jian

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Information)
in The University of Michigan
2010

Doctoral Committee:
Professor Jeffrey K. MacKie-Mason, Chair
Professor Yan Chen
Assistant Professor Nigel P. Melville
Assistant Professor Rahul Sami

We should be taught not to wait for inspiration to start a thing. Action always generates inspiration. Inspiration seldom generates action.

—Frank Tibolt

to Nisha

Acknowledgments

First and foremost, I am deeply indebted to my mentor, committee chair, and a caring friend, Jeffrey K. MacKie-Mason, who has been both uncompromising in the rigor of scientific research and limitlessly supportive in providing me valuable guidance. The experience of working with him will continue to shape the kind of scholar and teacher I will become in the years ahead. I thank my committee members, Yan Chen, Rahul Sami, and Nigel Melville, for providing me thoughtful comments and suggestions at various stages towards the completion of my dissertation.

Chapter 2 is based on my joint work with Jeffrey K. MacKie-Mason and Paul Resnick. Paul Resnick was invaluable for contributing an interesting research question and a large data set from eBay. This study, which took five years to complete, provided me with an opportunity to work with the best researchers in the field. The experience will constantly remind me to never settle with anything short of perfection. I also thank Tapan Khopkar, who participated in the early stage of this study and helped with data extraction, and Sijian Wang, who helped me verify the correctness of our statistical model. During this study I was financially supported in part by grant IIS-9977999 from the National Science Foundation awarded to Paul Resnick.

Chapter 3 is based on my joint work with Rahul Sami. I am grateful to Rahul Sami for his thoughtful contributions throughout the project. Matthew Bombyk and Roberto Jimenez provided excellent research assistance. This study was supported by grant CCF-0728768 from the National Science Foundation awarded to Rahul Sami and the research incentive funding from the School of Information provided also to Rahul Sami. A one-page extended abstract of this paper was published in the Proceedings of the 2010 ACM Conference on Electronic Commerce (EC10).

Chapter 4 was completed under the supervision of Jeffrey K. MacKie-Mason, who provided me with valuable comments and suggestions. I also thank Yan Chen, who pointed me to the area of excludable public goods, collaborated with me during the early stages of this study, and continued to support me in every opportunity possible despite her heavy workload in the past few years.

I thank the participants of the Incentive-Centered Design lab group and the Behavioral and Experimental Economics Lab at the School of Information, especially Rick Wash, Anna Osepayshvili, Tapan Khopkar, John Lin, Tracy Liu, Ben Stearns, and Josh Cherry. I would like to thank the School of Information for providing me with generous financial support for five years, and the Rackham Graduate School for awarding me the Predoctoral Fellowship, which supported the last year of my study.

Last but not least, I thank my husband Kabi, without the support of whom none of this would be possible. Thank you for believing in me and reminding me to continue with what I love to do, throughout all the ups and downs of my academic journey in the last six years. And of course, my precious little one, Nisha, has been a constant source of joy for me. I thank my other family members, Changzhen Yu, my mother, Xianfa Jian, my father, Ashalata Padhi, my mother-in-law, and Siba Prasad Padhi, my father-in-law for giving me so much support over the years, especially for helping me take care of my little one so that I could concentrate on completing my dissertation in the last year.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

Abstract

My dissertation contains three studies centering on the question: how to motivate people to contribute high quality information on information aggregation systems, also known as social computing systems? I take a social scientific approach to *identify* the strategic behavior of individuals in these systems, and *analyze* how non-monetary incentive schemes motivate information provision. In my first study, I use statistical modeling to infer users' information provision strategies from their actions. Information system users' strategies for contribution (e.g., I only contribute if others have contributed a certain amount) are often not directly observable, but identifying their strategies is useful in system design. With my co-authors, Jeffrey MacKie-Mason and Paul Resnick, I constructed a maximum likelihood model with simultaneous equations to estimate strategic feedback reciprocation (i.e., I only provide feedback if you give me feedback first) among the traders on eBay. We found about 23% of the traders strategically reciprocate feedback. My second study is focused on truthful provision of information in information markets — markets in which the participants trade bets about future events. The resulting market price reflects an aggregated prediction for the event. Theory predicts that when traders' private information is substitutable — contains similar information — they profit most by trading honestly. But when traders' private information is complementary — contains exclusively different information — traders are better off bluffing, i.e., first trading dishonestly to mislead others and later profiting from others' mistakes. Using human-subject experiments, my co-author Rahul Sami and I found traders bluff more in markets with complements than in markets with substitutes. In my third study, I use game theory to analyze two non-monetary mechanisms for motivating information provision: the minimum threshold mechanism (MTM), under which one can access the systems if she contributes more than a threshold, and the ratio mechanism (RM), under which a user consumes an amount proportional to her contribution level. I found whenever RM can achieve the social optimum, MTM can achieve the same. If RM implements a no-exclusion equilibrium, the same outcome can always be implemented by MTM.

Chapter 1
Introduction

Information systems can be promising tools for collecting and aggregating distributed information held by individuals. These systems, also known as social computing systems or Web 2.0, can provide valuable information for firms, communities, or society. For example, reputation systems support online transactions among strangers (e.g., peer-to-peer lending). Market-based information systems, e.g., prediction markets, are used widely to forecast future events in various settings to help with risk management, such as the sales of a firm, the likelihood of an influenza outbreak, or the enrollment of schools etc. Online communities such as Peer2Patent.org help the United States Patent and Trademark Office to identify crucial prior art in order to better examine patent applications.

My research is focused on the strategic behavior of individuals in information systems. While technology *enables* people to share information relatively inexpensively, people may or may not have the right incentives to do so. Humans are autonomous components of information systems. They have their own objectives that may or may not be aligned with what is intended by the information system designers. Contribution costs time and effort, which may discourage information provision. Some users of these systems may provide information that misleads others. In market-based information systems, imperfect information aggregation might occur due to traders manipulating the market prices.

My broad research question is: how can we motivate people to share high quality information? I take a social scientific approach to *understand* the strategic behavior of individuals in existing information systems, and I also *design* incentivizing mechanisms, and analyze and test how they can be effective in motivating high quality information provision.

To understand the strategic behavior of users of information systems is not always straightforward. While users' contribution actions (contribute or not) are directly observable, their strategies (e.g., I only contribute if others have contributed this much) are often more complex and not directly observable. However, users' strategies are often more informative than their observable actions to system managers in improving design. What factors influence the users' decisions to contribute and in what way? Are users' actions dependent on one another's actions? In chapter 2, I report my empirical study in which I use statistical methods to infer users' latent strategies from their observable actions. With my co-authors, Jeffrey MacKie-Mason and Paul Resnick, I studied the strategic provision of feedback in eBay's reputation system. This system relies on the buyers and sellers to *voluntarily* submit feedback about each other, so that it can publish each trader's reputation to deter dishonest behavior. Despite an apparent strong incentive to free-ride, more than half of the transactions actually receive feedback (Resnick and Zeckhauser, 2002). We conjecture that one possible reason for traders to provide feedback is because they engage in feedback reciprocation ("you scratch my back and I'll scratch yours"). We constructed a maximum likelihood model with simultaneous equations that allows us to estimate the prevalence of feedback reciprocation based on traders' observable actions, e.g., whether the feedback was provided and the time at which it was provided. Applying our model to a large data set containing about one million transactions on eBay, we found 20 to 23% of the traders on eBay are strategic feedback reciprocators: they only provide feedback if they receive feedback from their trading partners. We also further estimate the effect of multiple factors that affect eBay traders' feedback provision behavior, including both the buyer's and the seller's reputation profiles and the price of the item being traded.

In chapter 2, I have focused on whether the users of an information system provide their private information. In chapter 3, I turn to whether the users *accurately* provide their private information. I report a study on market-based information systems: prediction markets. In a prediction market, also known as an information market, the participants trade bets on the outcomes of a future event. The price of the market then reflects an aggregated prediction of the future event based on all the participants' private information. Such markets have been created for a wide range of applications; examples include the Iowa Electronic Market for forecasting elections and other political events,[1] the Hollywood Stock Exchange for forecasting movie box office receipts (HSX.com), and intra-company markets to forecast sales (Cowgill et al., 2008, Footnote 2).

The wide adoption of prediction markets does not mean we understand them well. Few factors that might influence the performance of prediction markets have

[1]See `http://www.biz.uiowa.edu/iem/index.cfm`, retrieved on July 25, 2010.

been studied. My co-author Rahul Sami and I identified three factors, either from prior literature or from questions raised by the designers of prediction markets. We then tested these three factors using human-subject laboratory experiments. The first factor we consider is in the choice of two commonly used trading mechanisms: a direct mechanism in which traders report their beliefs as probabilities, and an indirect mechanism in which traders reveal their beliefs through buying and selling securities. Theory suggests that as these two mechanisms are mathematically equivalent, the predictions generated by the two mechanisms should be equally accurate. Indeed, our experimental results indicate that the choice of either trading mechanism does not affect the accuracy of the market predictions significantly.

The second factor we study is in providing the structure of strictly sequenced opportunities to trade, as compared to the standard approach of letting traders choose when to trade in an unstructured way. We have two motivations in considering a structure: First, the existing theoretical results (Chen et al., 2007; Dimitrov and Sami, 2008; Chen et al., 2009; Dimitrov and Sami, 2010) implicitly assume a structured order of trading opportunities, and this experiment allows us to test if this assumption is of practical significance. Second, enforcing more structured interaction has been shown to help in group forecasting performance (Graefe and Armstrong, 2008). In our study, we found that structured markets generate significantly more accurate predictions than unstructured markets.

The third factor we study is manipulation that might occur in prediction markets. While one might assume that traders in these markets will truthfully reveal their beliefs in order to profit, such assumptions may or may not hold depending on the context in which the market is operated Chen et al. (2009). Chen et al. (2009) predict that when traders' private information are substitutes — traders' private signals are likely to be the same — they profit most when they trade according to their true belief and put in their trades as early as possible. But when the private information sets held by traders are complements — the information gain from knowing all the signals is greater than the sum of gains from each trader's signal— traders are better off delaying their trades or bluffing. Bluffing is a strategy that involves trading early to misrepresent their own private information in order to mislead other traders, then profiting from others' mistakes. We tested these predictions and found supporting evidence that traders delay their trades and bluff more when their private information are complements than when they are substitutes. In addition to the effect of the distribution of traders' private information, we also tested other factors that might influence the performance of prediction markets, namely the particular implementation of the market algorithm, and how much structure the market imposes on the traders' trading order. These tests generate practical implications for designing well-performing prediction markets, as I discuss in detail in chapter 3.

Both chapter 2 and 3 are about empirically identifying individual behaviors in specific information systems. In the last major chapter of my dissertation, chapter 4, I report a theoretical study that is aimed at generating specific design implications for encouraging information provision in future information aggregation systems. I impose two criteria on my design. First, it has to do at least as well as the most commonly used scheme — voluntary provision — and it should not rely on monetary payments from or to the users to motivate contribution, as for many systems, using micro-payments is not a practical solution for logistic or social reasons.

In chapter 4, I use game theory to analyze the performance of two simple non-monetary mechanisms: the minimum threshold mechanism, under which one can only access the public goods if her contribution is higher than a pre-specified threshold, and the ratio mechanism, under which a user consumes at most an amount proportional to her own contribution level. I derive equilibrium predictions for these two mechanisms and analyze their performance in terms of social welfare. My results indicate some advantages of the minimum threshold mechanism over the ratio mechanism. There exist some conditions under which the minimum threshold mechanism can achieve the social optimum, but the ratio mechanism cannot. Furthermore, if the ratio mechanism implements a no-exclusion equilibrium, the same outcome can always be implemented by the minimum threshold mechanism. I discuss the limitations of my study and the possible future directions in chapter 4.

These three studies are among the first steps towards systematically understanding strategic behavior of individuals in information aggregation systems. In chapter 5, I conclude and discuss some open questions related to my dissertation.

Chapter 2

I Scratched Yours: The Prevalence of Reciprocation in Feedback Provision on eBay

2.1 Introduction

Reputation systems enable trade among strangers by informing people about their trading partner's past performance, which also creates incentives for good behavior (Resnick et al., 2000; Dellarocas, 2003). Many online marketplaces offer reputation systems based on subjectively provided feedback. For example, buyers rate sellers on Amazon's marketplace, and those ratings are visible to future buyers.

In some cases, feedback provision is two-sided. For example, at couchsurfing.net, where travelers find free places to stay while traveling, both hosts and travelers can rate each other. eBay buyers and sellers can rate each other. Many other two-sided markets are also candidates for two-sided reputation systems, such as dating sites, housing matches, and ride-sharing services.

Subjectively reported two-sided feedback introduces strategic considerations: whether to provide feedback, and the content of that feedback, may be influenced by the partner's actual or expected feedback-giving behavior. For example, anecdotes suggest that some eBay users employ a feedback-giving strategy we call "reciprocation": they only give feedback after receiving feedback from their trading partners.[1] Anecdotal evidence also suggests that buyers and sellers withhold negative feedback

[1] For example, on Yahoo!Answers a user asked: "Why do eBay sellers not give feedback as soon as you pay?", and received the answer "Many sellers wait until they receive feedback from the buyer before they leave feedback" (`http://answers.yahoo.com/question/index?qid=20060816131534AAecgGz`, retrieved on Oct 28, 2008). Similar conversations also occur on eBay's forum (e.g., `http://reviews.ebay.com/Who-should-leave-FEEDBACK-first-BUYERS-or-SELLERS_W0QQugidZ10000000003772517`, retrieved on Oct 28, 2008).

in order to avoid receiving retaliatory negative feedback. eBay conducted experiments on alternative feedback designs and, in 2008, removed the option for sellers to provide negative feedback to buyers, though they can still provide positive feedback and buyers can still provide either kind (Bolton et al., 2009).

While two-sided subjective feedback may inhibit provision of negative feedback, it may help solve an underprovision problem for positive feedback. Feedback information is a public good: one person's consumption of published feedback does not diminish another's use of it. Theory predicts that in general public goods will be under-provided (Samuelson, 1954b). The free-rider problem seems especially pernicious because feedback only benefits other users, not its provider, so that self-interested users then appear to have little or no incentive to provide feedback. Nevertheless, more than half of the traders on eBay provide feedback (Resnick and Zeckhauser, 2002). Why?

There are many possible motivations, likely experienced to a greater or lesser degree by different people. For example, some may freely contribute feedback because they are altruists, willing to incur a small cost to contribute to the community (Fehr and Gächter, 2002). Some may exhibit "reciprocal altruism" (Andreoni and Miller, 1993; Gächter and Falk, 2002), a tendency to give people "what they deserve", in this case a positive feedback in return for a good transaction and a negative feedback in return for a bad one. Some may provide feedback to avoid the hassle of partners asking for or demanding it. Some may fear that if word gets around that they don't provide feedback, partners will take advantage of them (Gazzale, 2004, Chapter 1).

Another reason to provide feedback is that it may spark the desirable event of the partner providing feedback. If many of one's transaction partners employ a feedback reciprocation strategy, then providing feedback first can be a way to build one's own feedback profile faster. If everyone followed a strategy of reciprocating feedback, and no one chose to give unconditionally, no one would ever be the first to provide feedback and none would be provided. On the other hand, if no one followed a reciprocation strategy, one of the incentives for providing feedback would be removed. Estimating the prevalence of feedback reciprocation provides a window into the complex ecology of feedback provision.

Of course, that begs the question of why a self-interested party would employ a strategy of giving feedback after receiving it. Clearly, having one's partner expect such feedback reciprocation can induce the partner to provide positive feedback (in order to get it in return) and to remain silent rather than providing negative feedback after a bad transaction (in order to avoid getting it in return). In a one-shot game, however, in the absence of a binding feedback reciprocation contract, actually delivering the reciprocal feedback, at some cost of effort, would not be rational. With repeated interactions, some form of direct retaliation may be sufficient for a reciprocation equilibrium. Or, even without repeated direct interaction, a generalized reciprocity equilibrium may emerge (Jian and MacKie-Mason, 2008). As with any feedback provision, some reciprocators may also want to follow through on providing feedback for the non-rational reason of giving a gift to, or taking vengeance on their partners, in this case rewarding or punishing the partners' feedback rather than the partners' action in the underlying sales transaction.

We do not propose a theoretical model for why feedback reciprocation occurs: we would not be able to estimate a structural model with the data available to us in any case. Rather, using a large dataset of eBay transactions we test for the prevalence of reciprocation, and the prevalence of two alternative strategies, unconditional provision and non-provision. We find that buyers and sellers used the "reciprocate" strategy about 20-30% of the time. We also measure the extent to which the prevalence of these strategies changes with the experience levels of the two parties, and with the item price. For example, in bilateral transactions, the relative experience levels may matter as inexperienced traders learn the strategy equilibrium, while experienced traders may be trying to teach their partners.[2] Dellarocas and Wood (2008) have found that the level of traders' satisfaction varies with item price; we explore whether it affects the prevalence of reciprocal or non-reciprocal feedback giving strategies.

We also make methodological contributions to the estimation of feedback provision strategies in two-sided reputation systems. When both parties provided feedback, it is not clear whether they did so independently or whether the second did so in response to the first. When neither party provided feedback, it is not clear whether their decisions were unconditional or whether one or both would have provided feedback had the other done so. This is a problem in estimating choice when the underlying decision variables are latent (unobservable). We develop a latent variables estimation procedure that takes advantage of the observable timing of feedback provision to identify and estimate reciprocal feedback-giving strategies. We are not the first to develop econometric models to identify reciprocal feedback-giving strategies. Previously, Dellarocas and Wood (2008) developed a different model to study both the biases in the feedback ratings and the reciprocal feedback-giving behavior. Our models differ in various ways, as detailed in section 2.2.

[2] In another context, Wikipedia has a "welcoming committee" in charge of greeting new members, introducing the community's policies, guidelines, and social norms to them (Wikipedia, 2010).

2.2 Related Work

Previous work on feedback provision has estimated the impact of prior negative feedback in a seller's history on the buyers' willingness to provide negative feedback. Resnick and Zeckhauser hypothesized that buyers may "stone" sellers who have received negative feedback, becoming harsher in their assessments of later transactions (Resnick and Zeckhauser, 2002). Empirically, the probability of receiving a negative feedback goes up immediately following the receipt of one (Cabral and Hortaçsu, 2010). There are several possible explanations besides stoning, including slipping (the first and subsequent negatives were both the result of the same decline in seller quality) and slacking (the seller provides lower quality because of receiving the first negative). By modeling the different but overlapping time windows in which these different explanations would operate, Khopkar et al. (2005) showed that some of the effect is indeed due to stoning. This line of work does not address the role of feedback reciprocation between partners to a given transaction.

A few studies have provided empirical evidence for the *existence* of strategic feedback reciprocation by exploring the correlations between buyers' and sellers' feedback timing (Bolton et al., 2009; Dellarocas and Wood, 2008; Resnick and Zeckhauser, 2002), but none of these studies offers an estimate of the *prevalence* of strategic feedback reciprocation. Bolton et al. also conducted human-subject experiments to compare the effects of various feedback provision mechanisms on the efficiency of the electronic market.

Dellarocas and Wood (2008), hereafter DW, is closest in spirit to our analysis. DW have two main results: a *feedback bias* result — traders report different transaction outcomes (positive, neutral, and negative) with different probabilities, leading to biases in the aggregated probabilities of various outcomes — and a *feedback reciprocation* result — that feedback received increases the probability of feedback giving. Like DW, we identified the existence of feedback reciprocation among eBay traders. We go further and calculate the magnitude of reciprocation, and we also measure the extent to which some observable factors influence traders' choices to strategically reciprocate feedback.

The key difference between our models is that we make different assumptions on how the timing of the first feedback affects the likelihood that the receiver reciprocates. we assume the proportion of reciprocators in the population holds *constant* over the time at which they receive feedback from their trading partners, whereas in DW's model, by construction, the probability of a trader reciprocating *decreases* in the time the first feedback is given. For example, if we estimate that, for a certain type of item, 20% of buyers follow the reciprocate strategy, it means that, no matter when the seller gives a feedback, it always triggers 20% of the total buyers to give feedback, in addition to the buyers who were going to provide feedback anyway, but just hadn't done so yet. In DW's model, however, the number of buyers who are triggered to reciprocate is higher when the seller gives feedback on day 20 than when the seller does so on day 40.

Another major difference between our models is that our method does not require a parametric assumption on how the receipt of feedback affects the recipient's feedback timing distribution. DW estimate whether there is reciprocation by estimating whether receiving feedback increases the partner's subsequent hazard rate for the time-to-feedback distribution (Dellarocas and Wood, 2008, Section 4.1.2).[3] Rather than assuming a change in the hazard rate after a feedback is received, we explicitly model the prevalence of a strategy of reciprocating (i.e., giving feedback *because* they received it, when they would not have had the partner's feedback been withheld). We assume a distribution of time-to-feedback, and if we identify an increase in the total mass of the feedback distribution after a feedback is received, we interpret the excess mass as evidence of reciprocation.

One benefit of our approach is that we are then able to estimate how much various observable factors affect the *prevalence* of feedback reciprocation. In particular, we estimated how much the traders' prior feedback profiles, and the item price, affect the traders' feedback giving strategy choices. In principle, DW's model can be extended for similar analyses, though less straightforward. To obtain a qualitative result on how different observable factors affect traders' reciprocation behavior, one can estimate how they affect the magnitude of the changes in the hazard rate of the time-to-feedback distribution. To obtain a quantitative result like ours, since the prevalence of feedback reciprocation declines in the time of the first feedback, the result would have a flavor like this: " If the first feedback was given on day x, a y dollar increase in the item price leads to z percent increase in the probability that the recipient reciprocates."

[3]The hazard rate $h(t)$ of a failure time distribution $F(t)$ is defined as the rate of an event occurring given that it didn't happen in the past. That is, $h(t) = F'(t)/(1 - F(t))$.

2.3 Model Description

2.3.1 Feedback outcomes and strategies

We do not observe eBay traders' feedback provision strategies (that is, their internal mental plans), only their observable feedback-giving actions. All transactions on eBay result in one of the following five outcomes in terms of feedback provision: No Feedback, only the seller gives feedback (Seller Only), only the buyer gives feedback (Buyer Only), the seller gives feedback first and the buyer next (Seller First), or the buyer gives feedback first and the seller next (Buyer First). Any distribution of unobserved strategies would generate a distribution of these five observable outcomes. We develop an estimation strategy that allows us to econometrically identify the latent (unobservable) distribution of strategy choices, given the observable outcomes.

We posit that each trader adopts, on each transaction, one of three strategies: abstain from giving feedback (N), give unconditionally (Y), or reciprocate (R). Y means that the player gives feedback on the transaction regardless of whether the partner does. R means that the player gives feedback on that transaction only if, and only after, the partner does.

In our stochastic model, we express buyers' and sellers' strategies in probability terms. Let R_g denote the probability that a trader with role $R \in \{B, S\}$ plays strategy $g \in \{y, r, n\}$. For example B_y is the probability that a Buyer plays strategy Y of giving feedback unconditionally. Thus, we can think of (B_y, B_r, B_n) as a mixed strategy that a buyer will follow on a particular transaction. The mix may depend on the item price, the number of prior feedbacks each partner has received, and many unobserved and unmodeled characteristics of the transaction and the buyer.

Intuitively from Table 2.1 it is clearly not possible to identify statistically the prevalence of the three strategies merely from observing which of the five feedback outcomes is realized after a transaction. Immediately, from the first row, if no feedback was provided by either the seller or the buyer, we know that both chose to either abstain or reciprocate, but we cannot tell which. However, in some transactions we can identify the strategy from the outcome. For example, if only the seller (buyer) gave feedback (second and third rows), it must have been that the seller (buyer) chose the give unconditionally strategy, and the buyer (seller) choose to abstain. Last, if both the seller and the buyer gave feedback, with one first and the other later (fourth and fifth rows), we can infer the strategy of only one of them: the first giver must have chosen to give unconditionally, but the second could have chosen either an un-

Table 2.1: Mapping Strategies To Feedback Outcomes

Outcome	Buyer Strategy	Seller Strategy
No Feedback	N or R	N or R
Seller Only	N	Y
Buyer Only	Y	N
Seller First	Y or R	Y
Buyer First	Y	Y or R

conditional giving strategy Y (but happened to act slower than the first provider), or a reciprocation strategy R.

2.3.2 Feedback timing

Although we cannot uniquely identify strategies from outcomes alone, we have more information available to us: we have the time at which feedback was provided. In our dataset, for each feedback we observe the time at which it was given, denoted by t_s if given by the seller and t_b if by the buyer, expressed as offsets or time elapsed from the close of bidding on the auction of the item. This information is sufficient to enable us to identify the strategies. For example, notice that the probability of observing an outcome in which both provide feedback, but the buyer gives first, depends on the timing:

$$Pr(\text{Buyer First}) = B_y \cdot S_y \cdot Pr(t_s > t_b | B_y, S_y) + B_y S_r, \tag{2.1}$$

where t_s and t_b are the time at which the seller and the buyer give feedback respectively. Thus, if we could estimate $Pr(t_b < t_s | B_y, S_y)$, it would help in identifying the quantities B_y, S_y, and S_r.

Conditional on the trader's role, $r \in \{b, s\}$, and her strategy choice, $g \in \{y, r\}$, we assume her time of feedback, t_r, follows a distribution described by the probability density function $f_{rg}(t_r)$. For example, for a seller who plays the unconditional strategy Y, the probability that she gives feedback at time t_s is $f_{sy}(t_s)$. We assume the timing distributions of both buyers' and sellers' feedback are lognormal, and write the feedback timing distribution for sellers playing the Y strategy as $f_{sy}(t_s) = LNORM(t_s)$, and similarly, for buyers playing the Y strategy as $f_{by}(t_b) = LNORM(t_b)$.[4] We

[4]Section 2.4.2 and appendix A.1 explore alternative functional forms besides the lognormal distribution.

can obtain an unbiased estimate for f_{sy} (f_{by}) using only observations with the Seller (Buyer) Only outcome.

In order to provide intuitions on how we use feedback timing to separately identify unconditional feedback givers and reciprocators, we illustrate our model of feedback timing in Figure 2.1. Imagine a dataset containing lots of transactions with the buyer giving feedback first, all at time t_b. Some of the sellers who subsequently gave feedback were following an unconditional strategy Y, while other may have been following a reciprocation strategy R. For those following strategy Y, absent buyer feedback the probability density of seller feedback at time t_s is $f_{sy}(t_s)$. Even with the buyer feedback, among sellers following strategy Y, behavior before time t_b should follow the same distribution, and thus a fraction $\alpha = F_{sy}(t_b)$ should have given feedback before t_b, and only a fraction $(1 - \alpha)$ are left to give feedback after t_b. In other words, $Pr(t_s > t_b | B_y, S_y) = 1 - \alpha = 1 - F_{sy}(t_b)$. If there were no sellers following reciprocation strategy R, and we had estimates of B_y and S_y, the overall probability of the buyer giving feedback first, at time t_b, would be $B_y f_{by}(t_b) S_y (1 - \alpha)$. If the dataset shows that the probability of a seller giving feedback after t_b is greater than $S_y(1-\alpha)$, the extra must have come from the reciprocators, shown as the shaded area marked with S_r.

Figure 2.1 shows a specific case with Buyer First outcome, in which the buyer gives feedback at time t_b. In our dataset, t_b, as well as t_s, can vary across the whole time axis, e.g., $t_b^{'}$ in Figure 2.1, and we do not have the luxury of many transactions for each particular value of t_b. Using the same logic illustrated in Figure 2.1, however, any values of $LNORM$ parameters defining f_{sy} and f_{by} will determine a likelihood of each of the transaction observations, with their actual t_s and t_b when feedbacks are provided. Thus, maximum likelihood estimation can be used to select $LNORM$ parameters that best fit the observed data.

2.3.3 The likelihood function

With the definition of feedback outcomes, strategy space, and timing distribution function, we construct a multinomial maximum likelihood model with simultaneous equations. Let θ denote the vector of parameters to be estimated, which will be explained in Section 2.3.4. Equation 2.2 is the overall likelihood function of θ given all the observable response variables Z in our dataset. For each transaction i, Z_i consists of the feedback provision outcome m_i, $m_i \in M$ where $M = \{\text{No Feedback, Seller Only, Buyer Only, Seller First, Buyer First}\}$, and the

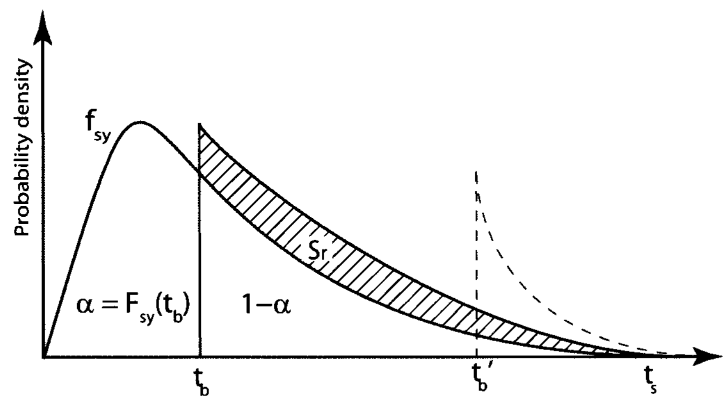

Figure 2.1: A simple version of the feedback timing model with the Buyer First outcome

times of those feedbacks, if they occur. Our likelihood function is as follows,

$$L(\theta; Z) = \prod_{i=1}^{N} l(\theta; m_i, t_b, t_s) \tag{2.2}$$

The outcome of No Feedback is observed whenever neither the seller nor the buyer played the strategy Y of giving feedback unconditionally. Thus, the likelihood of θ for No Feedback observations is,

$$l(\theta; \text{No Feedback}) = (1 - S_y)(1 - B_y), \tag{2.3}$$

When the outcome of Seller Only is observed, the likelihood is

$$l(\theta; \text{Seller Only}, t_s) = B_n S_y f_{sy}(t_s). \tag{2.4}$$

Similarly, when the outcome is Buyer Only, the likelihood is

$$l(\theta; \text{Buyer Only}, t_b) = S_n B_y f_{by}(t_b). \tag{2.5}$$

If the outcome of Buyer First is observed, the likelihood is as follows,

$$l(\theta;\text{Buyer First},t_b,t_s) = \tag{2.6}$$
$$\underbrace{B_y f_{by}(t_b)}_{\text{A}}\left(\underbrace{S_y}_{\text{B}}\left(\underbrace{Pr(t_s > t_b|B_y,S_y)}_{\text{C}}\right) + \underbrace{S_r}_{\text{D}}\right),$$

where term A contains the probability that the buyer was playing the strategy Y (give feedback unconditionally), and that he chooses this particular time, t_b, to give feedback; Terms B and D contain the probabilities that the seller might be playing the Y strategy and the R strategy respectively. Term C specifies the probability that if the seller is playing the Y strategy, her time of feedback happens to be later than the buyer's. Using F_{sy} to denote the corresponding cumulative distribution function of f_{sy}, we can write the Term C as:

$$Pr(t_s > t_b|B_y,S_y) = 1 - F_{sy}(t_b) \tag{2.7}$$

Similarly, if the outcome of Seller First is observed, the likelihood is as follows,

$$l(\theta;\text{Seller First},t_b,t_s) = S_y f_{sy}(t_s)\left(B_y Pr(t_b > t_s|B_y,S_y) + B_r\right), \tag{2.8}$$

where $Pr(t_b > t_s|B_y,S_y) = 1 - F_{by}(t_s)$.

2.3.4 Functional form assumptions

To estimate parametrically the probabilities of trader i playing each of the three strategies, we make the assumption that her probabilities of choosing any one of the three strategies are governed by multinomial logistic distributions:

$$\begin{aligned} B_{yi} &= \frac{e^{\beta_y X_i}}{1 + e^{\beta_y X_i} + e^{\beta_r X_i}}, \\ B_{ri} &= \frac{e^{\beta_r X_i}}{1 + e^{\beta_y X_i} + e^{\beta_r X_i}}, \\ B_{ni} &= \frac{1}{1 + e^{\beta_y X_i} + e^{\beta_r X_i}}, \end{aligned}$$

$$
\begin{aligned}
S_{yi} &= \frac{e^{\theta_y X_i}}{1+e^{\theta_y X_i}+e^{\theta_r X_i}}, \\
S_{ri} &= \frac{e^{\theta_r X_i}}{1+e^{\theta_y X_i}+e^{\theta_r X_i}}, \\
S_{ni} &= \frac{1}{1+e^{\theta_y X_i}+e^{\theta_r X_i}},
\end{aligned}
$$

where X_i is the vector of independent variables that we will later use in the regression.

2.3.5 Model validation

To validate the mathematical model and our STATA code, we conducted Monte Carlo simulations. We generated datasets according to our functional form assumptions and a set of arbitrarily chosen parameters. See appendix A.3 for the true values of all the parameters used in the simulation, and the estimates for these parameters found using our model. Summarizing our simulation results, the distribution of simulation errors is quite close to the predicted distribution. In the last column of Table A.3 we report the errors in units of standard deviations, and in Figure A.1 we plot the cumulative distribution of these errors against the asymptotic normal distribution the Monte Carlo should generate. The match is quite good, and there are no outliers. We conclude that our model is well identified and correctly programmed.

2.4 Data set

We derived our sample from three master datasets provided by eBay:

1. Items Dataset: contains transactional data for all the items listed for sale on eBay from February 1st 1999 to June 30th 1999.
2. Feedback Dataset: contains all feedback data up to May 31st 1999.
3. Users Dataset: contains the id and registration dates for all the users who registered before June 30th 1999.

Some buyer-seller pairs conduct multiple transactions, and feedback giving patterns may be quite different on subsequent transactions than on initial transactions, especially since eBay did not count multiple feedbacks from the same partner in a trader's score, thus potentially changing the incentive to provide multiple feedbacks to the same partner. Moreover, at the time of our dataset, eBay did not require that all feedback be tied to a specific transaction. Thus, we chose the buyer-seller partnership's first transaction as the unit of analysis.

We extracted all the items listed for auction during the first week of March 1999 and eventually purchased, involving buyer-seller partnerships that had not conducted a prior transaction and had no prior feedback. This initial sample contains 959,657 items.[5]

The auction of an item ends with a winning bidder, whose bid is higher than other bidders' bids and the reservation price set by the seller. From our perspective, this marks the start of a transaction.

At the time of these transactions, eBay opened the feedback channels for both buyer and seller to rate each other as soon as the transactions started. Each feedback contained two parts: an indicator (+1 for positive, -1 for negative, or 0 for neutral) and an optional text comment. We treated the first feedback, if any, from buyer to seller and vice versa, that occurred within 60 days of the transaction, as feedback for that first transaction between partners.[6]

2.4.1 Sampling

The two-sided nature of the feedback systems we study poses particular challenges for data sampling. The unit of analysis in our maximum likelihood model is a transaction, and we assumed above that feedback strategy selection for all the transactions in the dataset are pairwise independent, conditional on the item price and feedback profiles. Yet suppose that each trader (buyer or seller) has an idiosyncratic individual propensity to choose one of the three feedback strategies (always give, never give, reciprocate). Then if a trader participates in multiple transactions, these transactions will be inter-dependent, thus violating the independence assumption.

We present below a method that yields consistent estimates in the face of this multiple-transaction, fixed-effect problem. Before we do, we explain why a couple of other seemingly natural methods do not work. One is to randomly select one from all

[5] Our transaction data began only in February 1999, while our feedback history goes back to the beginning of eBay. There is a chance that we included transactions that were not the first for a partnership, if the previous transactions were more than one month prior to our extraction window and no feedback had ever been given between the partners.

[6] It is possible that some partnerships conducted additional transactions beyond the first and provided feedback for subsequent transactions but not the first. Our data set would incorrectly attribute the first feedback with the first transaction. We believe that occurred infrequently and would introduce only random error, not systematic bias.

transactions in which a given buyer participates, losing some observations but eliminating buyer interdependence. Unfortunately, since transactions involve trader pairs, this method will not, in general, eliminate all multiple transactions for some sellers. Further, if multiple-transaction buyers are matched approximately randomly to sellers, then this sampling method will disproportionately eliminate altogether sellers who trade infrequently, biasing the resulting sample of sellers. Sampling both sides to eliminate multiple transactions for both buyers and sellers simply exacerbates the second problem.

Another approach is common: to specify a maximum likelihood model that explicitly accounts for the possibility of fixed effects, estimating them as nuisance parameters, to yield consistent estimates of the parameters of interest. However, in our sample of 959,657 items, there are 394,997 distinct buyers, and 133,697 distinct sellers. Thus, on average buyers in our sample appear in about two transactions, and sellers appear in about seven. Indeed, 50% of the traders had no more than two transactions, and 90% had no more than nine transactions. For most traders, this longitudinal dimension is much too small to rely on asymptotics for consistency. Greene (2004) found that with such a small number of repeated observations per agent, both fixed and random effects models produce results that are more biased than those from a pooled model in which trader-specific effects are ignored.

We now describe our method. We constructed two sub-samples, one a **buyer-unique** sub-sample containing a single randomly drawn transaction for each of the 394,997 unique buyers, and the other a **seller-unique** sub-sample containing a single randomly drawn transaction for each of the 133,697 unique sellers. We then estimate the model twice, once for each sub-sample. We obtain estimates of the parameters associated with buyer behavior from the buyer-unique sub-sample (discarding the estimates of the seller parameters), and we obtain seller parameter estimates from the seller-unique sub-sample.

Our sampling method ensures a representative sample. The buyer (seller) sub-sample is a representative cross-section sample of the buyer (seller) population. More importantly, the sampling method ensures unbiased parameter estimates for the uniquely-sampled side, despite potential biases in the parameter estimates for the other side. For example, consider the seller-unique sub-sample. For any fixed buyer parameters, the likelihood function is maximized at the same set of seller parameters.[7] Thus, because the seller-unique sub-sample has independence among seller

[7]One can easily verify this by taking the partial derivative of the log-likelihood function with respect to any seller parameter: no buyer parameters appear. This property is due to our assump-

transactions, we get an unbiased estimate of the seller parameters, even if the buyer parameter estimates are not accurate due to dependence among transactions involving the same buyer.

This method may not be efficient, compared to a hypothetical panel model, because we discard many observations. A panel model with fixed or random effects, however, is not suited to our dataset, as argued above. Fortunately, with our rather large dataset, we can be somewhat profligate and still obtain rather precise estimates.

2.4.2 Model Fit

We tested three different parametric functional forms, i.e., Lognormal, Gamma, and Weibull distributions, to estimate f_{sy} and f_{by}. The observed distribution and the estimated distributions using these three functional forms are graphed in Figure 2.2 for sellers and Figure 2.3 for buyers.

From a visual comparison of the estimated and the observed distributions, we believe that the lognormal model fits the observed distributions better than the other two. However, the results of Kolmogorov-Smirnov tests reject the null hypotheses that the observed and the predicted timing distributions are the same, for all three functional forms and for both timing distributions, i.e., f_{sy} and f_{by}, each at a statistically significant level (p-value < 0.001). This is understandable because with a large dataset (more than one hundred thousand observations), almost any deviation from a hypothesis will be statistically significant. Thus the Kolmogorov-Smirnov test results are overly precise for determining the reasonableness of the goodness of fit of our model.

We also conducted robustness test of our results using all three distribution functions for feedback timing (See appendix A.1 for details). As varying the functional forms did not lead to qualitatively different results, we only report our results estimated using the lognormal distribution.

As a sanity check on the assumption that receiving feedback triggers reciprocation, Figure 2.4 and 2.5 plot the actual timing of feedback, for sellers and buyers respectively, after receipt of a feedback on day 15 and 35, as compared to the expected feedback if there were no reciprocators. Both show spikes in feedback giving the day or two immediately after receiving feedback. These spikes could be due, in part or in full, to a reminding effect. That is, people who would have given feedback

tion that in any transaction the buyer and the seller independently choose their feedback provision strategies.

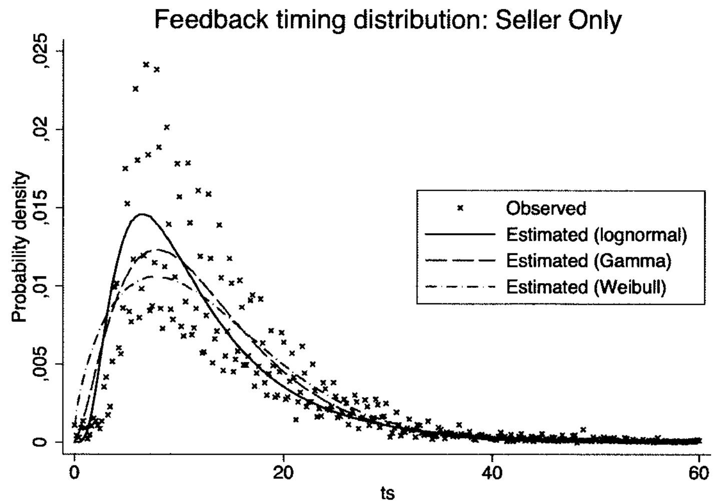

Figure 2.2: Sellers' feedback time distribution estimated using all the observations of Seller Only outcomes of our dataset.

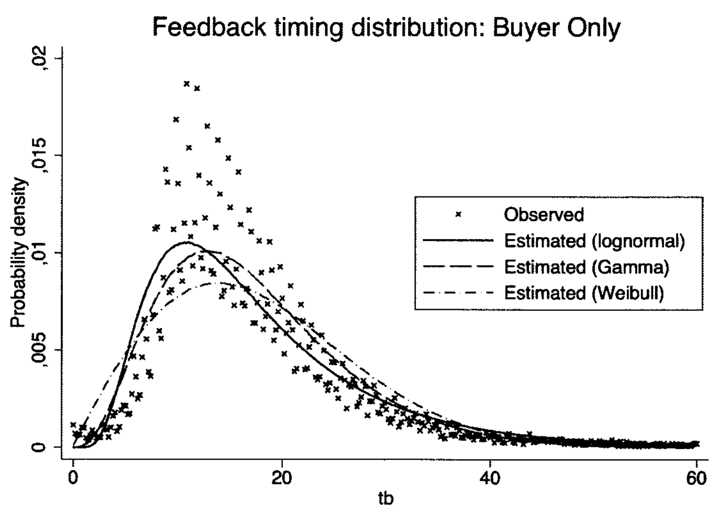

Figure 2.3: Buyers' feedback time distribution estimated using all the observations of Buyer Only outcomes of our dataset.

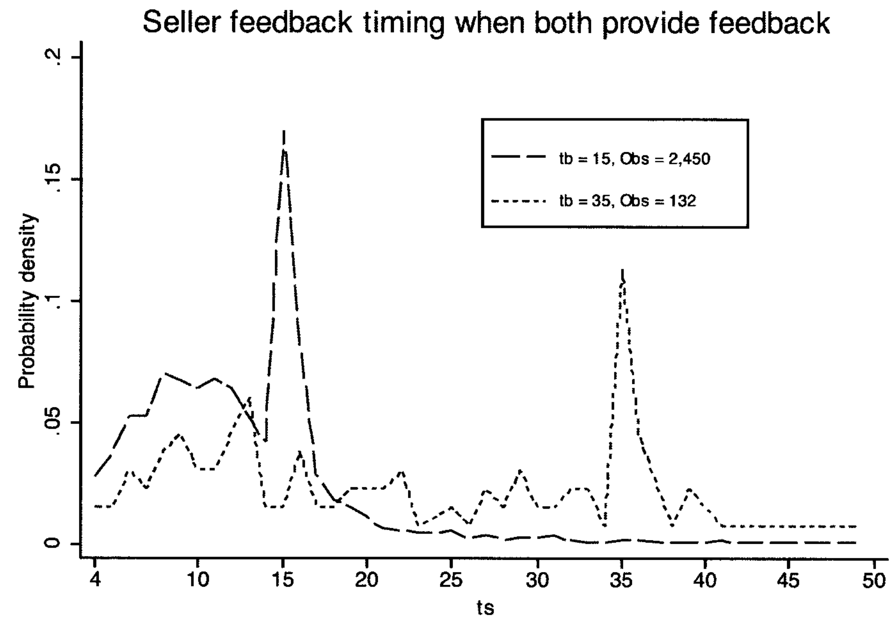

Figure 2.4: The distributions of seller feedback timing when buyers gave feedback on day 15 and day 35.

anyway may do so earlier, reminded by the event of receiving feedback. Reciprocation effects cannot be read off simply from these graphs: we estimate the reciprocation effect through our model, by comparing the area under the curve to the right of the first feedback event to the expected area, rather than examining the shape of those curves. The existence of the spike, however, provides clear evidence that receipt of the first feedback has some effect on the other party's feedback actions, and thus it is reasonable to attribute modeled changes in the second player's actions to the effect of the first party's feedback.

2.5 Empirical Results

Table 2.2 shows the distribution of feedback outcomes in both data sets: buyer-unique and seller-unique. Overall, about 35% of the transactions received no feedback; about 18% received feedback only from the seller; about 10% received feedback only from the buyer; and the remaining 35% received feedback from both parties. Table 2.3 shows the distribution of prior feedbacks received by the traders. Buyers have a mean of 26 feedbacks (with a standard error of 66), and a median of 6; sellers have a mean

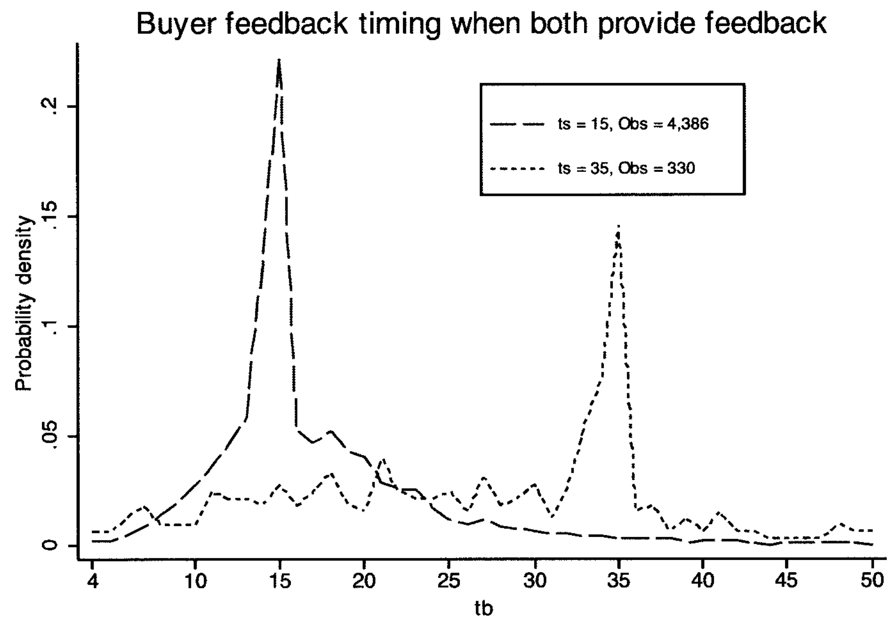

Figure 2.5: The distributions of buyer feedback timing when sellers gave feedback on day 15 and day 35.

	Buyer Unique		Seller Unique	
Outcomes	**# of Occurrences**	**%**	**# of Occurrences**	**%**
No Feedback	143,080	36.22%	45,927	34.35 %
Seller Only	73,276	18.55%	22,788	17.04 %
Buyer Only	40,669	10.3%	15,966	11.94 %
Seller First	90,564	22.92%	32,032	23.96 %
Buyer First	47,408	12%	16,984	12.71 %
Total	394,997	100%	133,697	100 %

Table 2.2: Distribution of feedback provision outcomes.

of 78 (with a standard error of 155), and a median of 26. Overall, sellers have more feedbacks than buyers. The distributions for both buyers and sellers are skewed: both distributions contain many traders with low feedback scores, and a few traders with high feedback scores. Note that when counting the total number of prior feedbacks for a trader, we followed eBay's practice (when it calculates feedback scores): we consider only the first feedback between a trading pair.

To study determinants of traders' choice of feedback provision strategies, we constructed six independent variables (see Table 2.4) to describe the observable context of the traders' transactions. We expect that the mix of feedback provision strategies

Roles	Mean(Std. Err.)	Median
Buyer	26(66)	6
Seller	78(155)	26

Table 2.3: Distribution of prior feedbacks received.

will depend on the reputation profiles of the participants, since at various stages of a trader's reputation profile development, her knowledge of the system varies, as well as her needs to maximize gains from her reputation profile. To elaborate, we first classify traders into two distinct categories: new and experienced traders. New traders are those who have very little experience with the feedback system, reflected in the small number (e.g., $0 \sim 4$) of feedbacks they have previously received. After participating in a few more transactions and receiving more than five feedbacks, we classify them as "experienced". The dummy variables *newbuyer* and *newseller* take the value 1 if a buyer or a seller is new, and zero otherwise. [8] *newint* is an interaction term of *newbuyer* and *newseller* to separately identify cases in which both the seller and the buyer are new. We also constructed two continuous variables to proxy for the traders' experience, $lnfbbuyerr$ and $lnfbsellerr$. They are the logarithm of the total number of prior feedbacks received in the trader's life time with eBay, for buyers and sellers respectively.[9] Last, we allow the logarithm of the price of the item being sold, *lnprice*, to be a factor in both the buyer's and the seller's strategy choices.

In Table 2.5 we present the coefficients on these independent variables estimated using our maximum likelihood model for four dependent variables: B_y, B_r, S_y, and S_r (B_n and S_n can be derived from these four variables). In the subsequent columns, we list their standard errors, z values, and the p-values of two-sided tests that they are not different from zero. Most estimates of the coefficients are significantly different from zero at 1% level. The coefficients in this table are hard to interpret, however, as they do not easily translate into marginal effects. In the following sections we evaluate the marginal effects of the independent variables, contingent on scenarios in which the independent variables take on various values.

[8]The choice of five as a threshold is arbitrary. We tested the sensitivity of our results to this threshold, and found that they *are* sensitive, but in the direction that reinforces our conclusion: newness matters, and experience effects show up after a modest number of feedbacks (see appendix A.2).

[9]To operationalize these two variables, we used $lnfbbuyerr = log(fbbuyerr + 1)$, and similarly for the seller $lnfbsellerr = log(fbsellerr + 1)$, to avoid the problem of an undefined logarithm of zero.

Variable name	Definition
newbuyer	1 if a buyer has received fewer than 5 feedbacks, and 0 otherwise
newseller	1 if a seller has received fewer than 5 feedbacks, and 0 otherwise
newint	an interaction term, equal to $newbuyer \times newseller$
lnfbbuyerr	the logarithm of the total number of feedback scores received by a buyer (*fbbuyerr*) plus one, thus $lnfbbuyerr = log(fbbuyerr + 1)$
lnfbsellerr	the logarithm of the total number of feedback scores received by a seller (*fbsellerr*) plus one, thus $lnfbsellerr = log(fbsellerr + 1)$
lnprice	log(the sale price of the item)

Table 2.4: Independent variables in the regression analyses.

2.5.1 Distribution of feedback provision strategies

In Table 2.6 we report our estimates of the probabilities of buyers or sellers playing any of the three hypothesized strategies, evaluated at the median and the mean number of feedbacks for each trader type. The top half of the table contains the estimated probabilities and the bottom half of the table contains the values of the independent variables at which these probabilities are evaluated. All the estimated probabilities shown in the table are statistically significantly different than zero at the 1% level. As we expected, a significant proportion of sellers and buyers are feedback reciprocators, and all three hypothesized strategies are being adopted for a substantial proportion of transactions. At the median levels (a buyer with a score of 6 buying a $45 item from a seller with score 26), there is a 38% probability the buyer will give feedback unconditionally and 39% probability she will abstain from giving feedback, with the remaining 23% probability she will be a reciprocator. On the other hand, at the median levels, 47% of sellers give feedback unconditionally; 32% abstain; and 20% reciprocate.

Depd Var.	Indp Var.	Coef.	Std. Err.	z	P-value
B_y	**newbuyer**	-0.113	0.019	-6.040	0.000
	newseller	-0.282	0.024	-11.780	0.000
	newint	-0.042	0.028	-1.500	0.134
	lnfbbuyerr	0.194	0.005	37.540	0.000
	lnfbsellerr	-0.03	0.003	-8.380	0.000
	lnprice	0.045	0.003	13.480	0.000
	intercept	-0.480	0.025	-19.200	0.000
B_r	**newbuyer**	0.18	0.036	4.910	0.000
	newseller	-0.286	0.046	-6.160	0.000
	newint	0.01	0.054	0.160	0.871
	lnfbbuyerr	0.098	0.010	9.670	0.000
	lnfbsellerr	-0.069	0.006	-11.640	0.000
	lnprice	-0.047	0.006	-7.290	0.000
	intercept	-0.310	0.048	-6.480	0.000
S_y	**newbuyer**	-0.280	0.036	-7.890	0.000
	newseller	-0.086	0.034	-2.540	0.011
	newint	-0.117	0.040	-2.950	0.003
	lnfbbuyerr	0.076	0.007	10.450	0.000
	lnfbsellerr	0.201	0.007	27.610	0.000
	lnprice	-0.045	0.006	-8.020	0.000
	intercept	-0.268	0.043	-6.190	0.000
S_r	**newbuyer**	-0.308	0.08	-4.000	0.000
	newseller	0.494	0.069	7.130	0.000
	newint	0.089	0.086	1.030	0.303
	lnfbbuyerr	0.030	0.014	2.070	0.039
	lnfbsellerr	0.204	0.015	13.870	0.000
	lnprice	0.102	0.012	8.700	0.000
	intercept	-1.60	0.092	-17.260	0.000

Table 2.5: Estimated coefficients.

Role and Strategy	Prob. at the median	Prob. at the mean
B_y	0.38 (0.002)***	0.43 (0.002)***
B_r	0.23 (0.003)***	0.22 (0.002)***
B_n	0.39 (0.003)***	0.35 (0.002)***
S_y	0.47 (0.004)***	0.52 (0.002)***
S_r	0.20 (0.005)***	0.21 (0.003)***
S_n	0.32 (0.005)***	0.26 (0.003)***
Independent Var. Values		
	$newbuyer = 0$	$newbuyer = 0$
	$newseller = 0$	$newseller = 0$
	$newint = 0$	$newint = 0$
	$fbbuyerr = 6$ (median)	$fbbuyerr = 26$ (mean)
	$fbsellerr = 26$ (median)	$fbsellerr = 78$ (mean)
	$price = 45.43$ (mean)	$price = 45.43$ (mean)

Table 2.6: Estimated probabilities of each strategy being adopted by typical sellers and buyers.

Note: Asterisks indicate the level of statistical significance. *** means statistically significant at the 1% level, ** at 5% and * at 10%. The numbers in parentheses are the standard deviations.

2.5.2 Strategy choices

Traders may choose different feedback provision strategies at different stages of their own career. They may also behave differently when facing different types of trading partners. In this section, we explore how traders' strategy choices vary based on the trading context.

New traders become experienced

We compare new sellers' and experienced sellers' strategy choices in Table 2.7. The variables with hats indicate that they are about the experienced sellers, and those without hats are about new sellers. We consider these comparisons in two scenarios, one in which sellers face new buyers ($newbuyer = 1$, $fbbuyerr = 0$), and the other in which they face experienced buyers ($newbuyer = 0, fbbuyerr = 26$, where 26 is the mean of $fbbuyerr$). The results are quite similar in these two scenarios (compare across the two columns in Table 2.7), so we focus on one: sellers facing new buyers (the first column of results). In exactly the same format, Table 2.8 shows the comparisons of new buyers' and experienced buyers' strategy choices. Again we focus on the first column of the results in Table 2.8.

Our results show that the changes in strategic behavior among sellers and buyers as they gain their first few feedbacks follow the same pattern: a reduction in the use of the "abstain" and "reciprocate" strategies, accompanied by a strong increase in "give unconditionally". The effects are larger in absolute and relative terms for sellers.

As new traders become experienced, they are more likely to give feedback unconditionally ($\widehat{S_y} - S_y = 0.12$, one-sided test: p-value < 0.001; and $\widehat{B_y} - B_y = 0.1$, one-sided test: p-value < 0.001), less likely to abstain from giving feedback ($\widehat{S_n} - S_n = -0.06$ one-sided test: p-value < 0.001; $\widehat{B_n} - B_n = -0.06$ (one-sided test: p-value < 0.001), and less likely to reciprocate ($\widehat{S_r} - S_r = -0.06$, one-sided test: p-value < 0.001; $\widehat{B_r} - B_r = -0.04$, one-sided test: p-value < 0.001). The probability of sellers (buyers) giving feedback unconditionally increases by 0.12 (0.1) point, which is a 55% (43%) increase from the new sellers' (buyers') probability of giving feedback unconditionally, $S_y = 0.22$ ($B_y = 0.23$), indicating that there is considerable learning going on among new traders during their first few trades. eBay provides FAQs and forums to facilitate learning. Also, the socially interactive nature of the feedback system makes it possible for new traders to learn by doing. Receiving feedback from one's trading partner informs or reminds the trader about the existence of the feedback system, and with it comes a link which guides her to return a feedback to her partner. In addition, receiving a feedback informs one about the social norm of feedback giving. Humans have the natural tendency to conform to social norms (Asch, 1956; Akerlof, 1980; Bernheim, 1994). All of these reasons point to the tendency that as new traders gain experience, they are more likely to be unconditional-givers, and less likely to abstain from giving feedback. We do not have a strong conjecture as to whether new traders will be more likely to reciprocate as they become experienced. The results show that they will not. One possible explanation is that before new traders learned how to use the feedback system, some of them were "passive" reciprocators who did not know how to give feedback but were willing to reciprocate any feedback received. Some traders who had learned how to give feedback and also that others may reciprocate their feedback, started actively initiating feedback exchanges. Thus these reciprocators have instead become unconditional givers.

How experienced traders treat new traders

In our dataset, 180,433 (46%) out of the 394,997 buyers are new buyers, and 31,609 (24%) of the 133,697 sellers are new sellers. If the social nature of the feedback system

Seller Strategy Prob.	Facing new buyer	Facing experienced buyer
$\widehat{S_y}$	0.35(0.004)***	0.45(0.004)***
$\widehat{S_r}$	0.16(0.007)***	0.18(0.005)***
$\widehat{S_n}$	0.5(0.007)***	0.38(0.005)***
S_y	0.22(0.004)***	0.33(0.004)***
S_r	0.22(0.01)***	0.24(0.006)***
S_n	0.56(0.01)***	0.43(0.005)***
$\widehat{S_y} - S_y$	0.12(0.006)***	0.12(0.00[illegible])***
$\widehat{S_r} - S_r$	−0.06(0.012)***	−0.06(0.007)***
$\widehat{S_n} - S_n$	−0.06(0.012)***	−0.06(0.007)***
Independent Var. Values		
	$newseller = 1$	$newseller = 1$
	$fbsellerr = 0$	$fbsellerr = 0$
	$\widehat{newseller} = 0$	$\widehat{newseller} = 0$
	$\widehat{fbsellerr} = 5$	$\widehat{fbsellerr} = 5$
	$newbuyer = 1$	$newbuyer = 0$
	$fbbuyerr = 0$	$fbbuyerr = 26$ (mean)
	$price = 45.43$ (mean)	$price = 45.43$ (mean)

Table 2.7: Seller behavior changes when new sellers become experienced.

plays a significant role in assisting new traders to learn, the speed at which they learn apparently depends on how soon they receive feedback. From the system designer's point of view, if the new traders received "special treatment" by veterans, they might learn faster. We have observed such special treatment in other communities, such as Wikipedia (Wikipedia, 2010). On eBay, do experienced traders also take up the responsibility of teaching new members? If so, they would be more likely to give feedback unconditionally to new traders than to other veterans.

Table 2.9 and Table 2.10 show how traders' behavior changes when their partners' reputation profiles vary. Again, variables with hats indicate that they are for experienced traders. For instance, $S_y(\widehat{B})$ denotes the probability of a seller giving feedback unconditionally when facing an experienced buyer, and $S_r(B)$ denotes the probability of a seller playing the reciprocate strategy when facing a new buyer.

The results do not bear out our conjecture that there may be some "indoctrination" going on among the traders. We found experienced traders do not educate newbies by giving them more feedbacks; rather, they give newbies fewer. Both sellers and buyers are more likely to give feedback unconditionally to experienced traders than to newbies: $S_y(\widehat{B}) - S_y(B) = 0.07$ (one-sided test: p-value < 0.001)

Seller Strategy Prob.	Facing new seller	Facing experienced seller
$\widehat{B_y}$	0.34(0.004)***	0.37(0.002)***
$\widehat{B_r}$	0.24(0.007)***	0.22(0.003)***
$\widehat{B_n}$	0.43(0.006)***	0.41(0.003)***
B_y	0.23(0.003)***	0.27(0.002)***
B_r	0.27(0.007)***	0.26(0.003)***
B_n	0.49(0.006)***	0.47(0.002)***
$\widehat{B_y} - B_y$	0.1(0.005)***	0.1(0.003)***
$\widehat{B_r} - B_r$	−0.04(0.009)***	−0.04(0.004)***
$\widehat{B_n} - B_n$	−0.06(0.009)***	−0.06(0.003)***
Independent Var. Values		
	$newbuyer = 1$	$newbuyer = 1$
	$fbbuyerr = 0$	$fbbuyerr = 0$
	$\widehat{newbuyer} = 0$	$\widehat{newbuyer} = 0$
	$\widehat{fbbuyerr} = 5$	$\widehat{fbbuyerr} = 5$
	$newseller = 1$	$newseller = 0$
	$fbsellerr = 0$	$fbsellerr = 78$ (mean)
	$price = 45.43$ (mean)	$price = 45.43$ (mean)

Table 2.8: Buyer behavior changes when new buyers become experienced.

and $B_y(\widehat{S}) - B_y(S) = 0.04$ (one-sided test: p-value < 0.001); they are less likely to abstain when trading with experienced traders: $S_n(\widehat{B}) - S_n(B) = -0.08$ (one-sided test: p-value < 0.001) and $B_n(\widehat{S}) - B_n(S) = -0.05$ (one-sided test: p-value < 0.001). One possible explanation is that for a considerable proportion of experienced traders, the purpose of giving feedback unconditionally is to initiate feedback exchanges. As new members are less familiar with the system, they may be less likely to "return the favor". As a result, experienced traders do not have sufficient incentives to give feedback to new traders.

Both veteran sellers and buyers are slightly more likely to reciprocate as their partners become experienced: $S_r(\widehat{B}) - S_r(B) = 0.02$ (one-sided test: p-value < 0.036) and $B_r(\widehat{S}) - B_r(S) = 0.007$ (not statistically significantly different from zero, but the change is in the right direction).

Gaining experience

Once an experienced seller has accumulated a substantial number of feedbacks, the marginal benefit of receiving a positive feedback diminishes, while the damage caused

Variable	Estimates
$S_y(\widehat{B})$	0.50(0.004)***
$S_r(\widehat{B})$	0.22(0.005)***
$S_n(\widehat{B})$	0.28(0.005)***
$S_y(B)$	0.43(0.004)***
$S_r(B)$	0.20(0.007)***
$S_n(B)$	0.37(0.006)***
$S_y(\widehat{B}) - S_y(B)$	0.07(0.006)***
$S_r(\widehat{B}) - S_r(B)$	0.02(0.009)**
$S_n(\widehat{B}) - S_n(B)$	−0.08(0.007)***
Independent Var. Values	
	$newbuyer = 1$
	$fbbuyerr = 0$
	$\widehat{newbuyer} = 0$
	$\widehat{fbbuyerr} = 5$
	$newseller = 0$
	$fbsellerr = 78$ (mean)
	$price = 45.43$ (mean)

Table 2.9: Seller behavior changes when new buyers become experienced.

Variable	Estimates
$B_y(\widehat{S})$	0.43(0.003)***
$B_r(\widehat{S})$	0.24(0.004)***
$B_n(\widehat{S})$	0.33(0.003)***
$B_y(S)$	0.41(0.003)***
$B_r(S)$	0.22(0.004)***
$B_n(S)$	0.38(0.003)***
$B_y(\widehat{S}) - B_y(S)$	0.04(0.005)***
$B_r(\widehat{S}) - B_r(S)$	0.007(0.007)
$B_n(\widehat{S}) - B_n(S)$	−0.05(0.006)***
Independent Var. Values	
	$newseller = 1$
	$fbsellerr = 0$
	$\widehat{newseller} = 0$
	$\widehat{fbsellerr} = 5$
	$newbuyerer = 0$
	$fbbuyererr = 26$ (mean)
	$price = 45.43$ (mean)

Table 2.10: Buyer behavior changes when new sellers become experienced.

by a negative feedback becomes salient. To avoid receiving negative feedback, sellers may choose never to give feedback first, which means she would either reciprocate or abstain. In repeated interactions with the same buyers or in situations in which the seller's past feedback giving behavior with other buyers is observable, establishing a reputation as a reciprocator rather than an abstainer would help the seller create some reward and retaliation power. Taken together, we expect sellers with higher number of feedbacks to be more likely to reciprocate and less likely to give feedback unconditionally. For the buyers, we do not have a strong conjecture as to how they will behave differently as they gain experience.

In Table 2.11 we report the marginal effects on feedback giving evaluated at the sample medians and means of $fbbuyerr$ and $fbsellerr$. Each marginal effect is reported in terms of probability on the scale from 0 to 100 percentage points. For example, $fbsellerr\_on\_Sy$ — the effect of the total number of prior feedbacks received by the seller on her probability of playing strategy Y — is 0.11 percentage points with a standard deviation of 0.01, which reads: as a seller receives one more feedback, the probability that she gives feedback unconditionally increases by 0.11 percentage points, all else being equal.

As expected, we found sellers are more likely to reciprocate when they gain experience: $fbsellerr\_on\_Sr = 0.05$ (one-sided test: p-value < 0.001) at the median level and $fbsellerr\_on\_Sr = 0.01$ (one-sided test: p-value < 0.001) at the mean level. Thus, for a seller with a median (mean) number of feedbacks, adding 100 more feedbacks increases the probability that he reciprocates by 5 (1) percentage points. To our surprise, we found sellers are more likely to give feedback unconditionally as they gain experience ($fbsellerr\_on\_Sy = 0.11$, one-sided test: p-value < 0.001). That is, for every 100 feedbacks a seller receives, she is 11 percentage points more likely to give feedback unconditionally. This effect is smaller when evaluated at the mean level, but still significant: $fbsellerr\_on\_Sy = 0.03$, one-sided test: p-value < 0.001.[10] These increases in S_y and S_r come from a reduction in S_n: $fbsellerr\_on\_Sn = -0.16$ (or -0.05 evaluated at the means). Taken together, as experienced sellers receive even more feedback, they are more likely to give feedback unconditionally or reciprocate, and less likely to abstain from giving feedback.

Turning to the effect of increasing feedback on experienced buyers, we found buyers are more likely to give feedback unconditionally as they gain more experience:

[10] We do not mean to claim that the estimated marginal effect holds constant over a range of experience from 26 to 126 feedbacks; rather, we are simply rescaling the coefficient magnitude to improve comprehension of the effects.

$fbbuyerr\_on\_By = 0.53$ at the median level (one-sided test: p-value < 0.001) and $fbbuyerr\_on\_By = 0.14$ at the mean level (one-sided test: p-value < 0.001). That is, buyers with 16 rather than 6 feedbacks go from 38% to 43.3% probability of choosing strategy Y of giving feedback unconditionally.[11] At the mean levels, buyers with 36 rather than 26 feedback go from 43% to 44.4% probability of choosing strategy Y.

How experienced traders are treated

We expect that experienced traders receive varying treatments based on the number of feedbacks they have accumulated. For example, as experienced buyers gather more (positive) feedback, they may seem more trustworthy as feedback givers. We expect that sellers are more willing to initiate a feedback exchange with trustworthy buyers.

As for experienced sellers, the more feedbacks they have, we would expect the lower the probability that they will receive feedback. Suppose seller A has 2000 feedbacks, while seller B has 100. Although both are experienced by our definition, we suspect that a buyer may be more likely to give feedback to B than to A, the reason being that the buyer feels her "vote" counts more for B than for A.

We did find that sellers are more likely to give feedback to buyers with more feedback: $fbbuyerr\_on\_Sy = 0.23$ (one-sided test: p-value < 0.001) at the median level, and $fbbuyerr\_on\_Sy = 0.06$ (one-sided test: p-value < 0.001) at the mean level. That is, for 10 new feedbacks a typical buyer receives, the probability that her partner seller gives feedback unconditionally to her increases by 2.3 percentage points at the median level, and by 0.6 percentage points at the mean level. This increase is mainly a decrease in S_n. Taken together, increasingly experienced buyers are more likely to always receive feedback from sellers. This result is consistent with our earlier results on how sellers treat buyers when we divide buyers into two groups: new versus experienced.

We did not find strong evidence to support our conjecture that buyers are less likely to give feedback unconditionally as sellers accumulate more experience: $fbsellerr\_on\_By = 0.00$. We did find (evaluating at the medians) that as sellers accumulate more feedback, buyers are less likely to reciprocate and more likely to abstain from giving them feedback: $fbsellerr\_on\_Br = -0.04$ (one-sided test: p-value < 0.001) and $fbsellerr\_on\_Bn = 0.04$ (one-sided test: p-value < 0.001).[12]. It appears

[11] Again, this is not quite true, since the marginal effect does not hold constant over the range from 6 to 16 feedbacks.

[12] These results are consistent with Dellarocas and Wood (2008): buyers are more likely to give feedback to inexperienced sellers than to experienced sellers.

Marginal effect	Evaluated at median	Evaluated at mean
$fbsellerr\_on\_Sy$	$0.11(0.01)^{***}$	$0.03(<0.00)^{***}$
$fbsellerr\_on\_Sr$	$0.05(0.01)^{***}$	$0.01(<0.00)^{***}$
$fbsellerr\_on\_Sn$	$-0.16(0.01)^{***}$	$-0.05(<0.00)^{***}$
$fbbuyerr\_on\_By$	$0.53(0.02)^{***}$	$0.14(<0.00)^{***}$
$fbbuyerr\_on\_Br$	$0.01(0.02)$	$-0.01(0.01)$
$fbbuyerr\_on\_Bn$	$-0.54(0.02)^{***}$	$-0.14(<0.00)^{***}$
$fbbuyerr\_on\_Sy$	$0.23(0.02)^{***}$	$0.06(0.01)^{***}$
$fbbuyerr\_on\_Sr$	$-0.03(0.03)$	$-0.01(0.01)^{*}$
$fbbuyerr\_on\_Sn$	$-0.19(0.03)^{***}$	$-0.04(0.01)^{***}$
$fbsellerr\_on\_By$	$-0.00(<0.00)$	$-0.00(<0.00)$
$fbsellerr\_on\_Br$	$-0.04(<0.00)^{***}$	$-0.01(<0.00)^{***}$
$fbsellerr\_on\_Bn$	$0.04(<0.00)^{***}$	$0.01(<0.00)^{***}$
$price\_on\_Sy$	$-0.05(<0.00)^{***}$	$-0.05(<0.00)^{***}$
$price\_on\_Sr$	$0.05(<0.00)^{***}$	$0.05(<0.00)^{***}$
$price\_on\_Sn$	$0.00(<0.00)$	$-0.00(<0.00)$
$price\_on\_By$	$0.03(<0.00)^{***}$	$0.03(<0.00)^{***}$
$price\_on\_Br$	$-0.03(<0.00)^{***}$	$-0.03(<0.00)^{***}$
$price\_on\_Bn$	$-0.01(<0.00)^{***}$	$-0.01(<0.00)^{***}$
Independent Var. Values		
	$newbuyer = 0$	$newbuyer = 0$
	$newseller = 0$	$newseller = 0$
	$fbbuyerr = 6$ (median)	$fbbuyerr = 26$ (mean)
	$fbsellerr = 26$ (median)	$fbsellerr = 78$ (mean)
	$price = 45.43$ (mean)	$price = 45.43$ (mean)

Table 2.11: Marginal effects on feedback giving evaluated at the sample medians and means of the number of feedbacks received by buyers and sellers.

our conjecture is in the right direction: buyers stopped bothering about giving feedback (or returning feedback) to highly experienced sellers, but the effects are quite small.

Item value

Items with different values may spark different feedback behavior. We know that buyers pay more attention to their sellers' feedback profile when buying higher valued items (Ba and Pavlou, 2002). We suspect that buyers are more likely to give feedback unconditionally if the prices of the items are higher. Anticipating buyers' behavior, sellers may be safer to strategically reciprocate rather than give feedback first, to avoid negative feedback.

We did find that item value affects the strategy choices of both sellers and buyers. Buyers are more likely to give feedback unconditionally when the price of the item is higher: $price\_on\_By = 0.03$ (one-sided test: p-value < 0.001). We also found with high value items, buyers are less likely to reciprocate: $price\_on\_Br = -0.03$ (one-sided test: p-value < 0.001). Thus for a \$100 increase in the price of the item, the buyers are three percentage points more likely to give feedback unconditionally, and three percentage points less likely to reciprocate. On the seller side, we found sellers are more likely to reciprocate, and less likely to give feedback unconditionally with higher value items: $price\_on\_Sr = 0.05$ (one-sided test: p-value < 0.001), and $price\_on\_Sy = -0.05$ (one-sided test: p-value < 0.001). Thus with a \$100 increase in the item value, sellers are five percentage points more likely to be reciprocators, and five percentage points less likely to give feedback unconditionally. These results are consistent with findings from prior literature (Ba and Pavlou, 2002; Dellarocas and Wood, 2008): buyers are more likely to give feedback for higher value items, and they are pickier in assessing the quality of the services for higher value items; in response to buyers' behavior, sellers tend to strategically hold back their feedback to retain the option to give retaliating negative feedbacks.

2.6 Limitations

In our specification we assumed that buyers and sellers choose their feedback giving *strategies* independently, conditional on item price and feedback profiles. This does not mean that they select actions independently: in particular when someone chooses the reciprocation strategy, the action of giving or not giving feedback depends on whether the partner does. Independent strategy choice does mean that buyer and seller did not collude (e.g., by making an outside agreement) when selecting their strategies of whether to give feedback conditionally or unconditionally. This is a common assumption for online transactions, and is unlikely to be problematic in the case of most eBay transactions.

There is another implication of our strategy independence assumption that may be of greater concern. Our strategy independence assumption is an assumption that strategy selection is not conditional on variables unobserved by the econometrician, but observable (at least in part) by both parties.[13] We may expect that strategies in fact depend on such variables. For instance, a person's decision whether to send feedback might depend on the other party's timeliness in carrying out his or her part of the transaction.

The problem can be illustrated by referring back to one of the terms that enters our likelihood function; consider the likelihood given that the outcome of No Feedback is observed:

$$l(\theta; \text{No Feedback}) = (1 - S_y)(1 - B_y), \tag{2.3}$$

That the strategies S_y and B_y are chosen independently absent collusion is not problematic. However, real users — say, sellers — might choose a "give feedback" strategy S_y that depends on the quality of the buyer's performance (as in, "give feedback always, if the buyer sends a check within three days") or of communications between the buyer and seller (as in, "give feedback always, if the buyer announces that he will give feedback"), and if the buyer's strategy also depends on some of the same variables, then the actions chosen by S_y and B_y may be correlated, and our likelihood function is misspecified.

The problem of omitting correlated variables on which strategies are contingent is a fundamental identification problem for all latent variable models, and is not special to our dataset nor our specification. The empirical question is how good the conditional independence assumption is (that is, that the econometrician observes and conditions on all salient correlating variables). Studies of other datasets, with different or more conditioning variables, would be valuable to test the robustness of our main claims. All that we can say, as is usual, is that our results are conditional on

[13]We use "observable in part" to refer to the case in which the two parties may observe "different" variables, but these variables themselves have some common component, or a joint non-independent distribution. Then, the correlated part of these two variables can lead to correlation in the actions chosen.

the specification, which in our case means that we assume strategy choices by both buyer and seller are not (very much) conditional on unobservables.

Another limitation of our study pertains to our dataset. The transactions in our dataset occurred in 1999. eBay has revised its feedback rules multiple times over the last decade since our data collection.[14] A major change occurred in May 2008, when eBay removed sellers' ability to leave negative or neutral feedback on buyers, to free dissatisfied buyers from fear of retaliation when they leave negative feedback.[15] Under this new rule, sellers lost their power of retaliation, which was one of the reasons *why* sellers may have strategically reciprocated feedback during the time our data were collected. We do not know how such rule revisions may have affected the population prevalence of eBay trader feedback provision strategies. Our methodological contribution provides a straightforward way to measure the effect of a rule change on feedback provision strategies using a before and after dataset.

In May 2007, eBay introduced Detailed Seller Ratings (DSRs) which enabled buyers to provide feedback on "four aspects of their transaction: accuracy of item description, communication, shipping time, and shipping and handling charges. The rating system is based on a one to five star scale, with one star being the lowest rating and five stars being the highest." The average ratings of a seller on all four aspects are displayed as part of her feedback profile.[16] The current system also displays the percentage of positive feedback out of the total number of positive and negative feedbacks received in the last 12 months.[17] These changes certainly enrich the display of the reputation profile, but the underlying mechanism remains largely the same. As long as the number of positive, neutral, and negative feedbacks are displayed, traders continue to care about receipt of these feedbacks when they formulate their feedback giving strategies.

eBay has also revised its feedback removal rules. At the time of our data collection, eBay did not allow revising feedbacks unless there was clear indication of feedback abuse. The current policy is that buyers can revise their negative or neutral feedback

[14]See eBay's official archive for details of these revisions at `http://www2.ebay.com/aw/au/archive.shtml`, retrieved on Mar 31, 2009.

[15]A number of measures were subsequently taken to protect sellers from buyers abusing their power conferred by this rule change, including enabling the buyer to revise her negative or neutral feedback if the seller managed to rectify the transaction problem. eBay also made a few other changes in May 2008, such as reducing the 90 day window of feedback giving to 60 days.

[16]See more details about DSRs at `http://pages.ebay.com/help/announcement/25.html`, retrieved on Mar 30, 2009.

[17]When the percentage of positive feedback was first introduced in March 2003, it was calculated as the percentage of positives out of the total number of positive, neutral, and negative feedbacks. In July 2008, eBay removed neutrals from the calculation.

if the seller manages to rectify the problem that led to the feedback. Such a policy change might affect buyers' attitudes toward giving negative or neutral feedback. We did not study the content of feedback, just strategies for whether to provide feedback. As we discussed earlier, however, the opportunity to retaliate, and now to revise, may affect traders' choice of feedback provision strategies as well.

The evolution of the rules in eBay's feedback system reflects the fact that these rules are important: the system managers consider it important enough to have adjusted these rules multiple times to achieve better feedback outcomes. It also implies that traders do respond to these rules. Given the rule changes, some of our results may no longer hold for the current eBay feedback system. Nonetheless, our results continue to provide a baseline estimate of the prevalence of the three feedback giving strategies, for comparison with other feedback systems in other electronic markets, including today's eBay. And perhaps equally important, we have developed a straightforward method that can be used to identify the prevalence of feedback giving strategies in other data sets, under varying environments. Because the feedback ecosystem is complex, no single study can account for the effects of future changes in the reputation system. Thus, having a reusable estimation method is valuable.

2.7 Conclusions

We developed an econometric model to study the feedback provision strategies used by participants in systems for bilateral interactions between strangers. We then applied our model to analyze the feedback provision strategies of eBay traders. We hypothesized that three types of feedback provision strategies were played by the traders: give (feedback) unconditionally, abstain (from giving feedback) unconditionally, and reciprocate. We found that all three types of strategies were being played by the traders. In particular, in transactions in which the buyer has the median number of feedbacks among all buyers, and the seller has the median number of feedbacks among all sellers, 38% of the buyers and 47% of the sellers give feedback unconditionally. This is quite high compared to theoretical predictions that the proportion of traders (buyers or sellers) who give feedback unconditionally would be minimal. We also found that in a substantial faction of cases, traders were strategic feedback reciprocators— 23% of the buyers and 20% of the sellers. We argue that the knowledge about the existence of these reciprocators may be a motivation for some traders to give feedback unconditionally, as they anticipate their partners to reciprocate. The remaining 39% of the time for buyers and 32% for sellers, the chosen strategy is not to provide feedback, regardless of whether the partner does.

eBay traders' feedback provision strategies evolve as they participate more in the marketplace. Both new buyers and new sellers become more likely to give feedback unconditionally after they experience their first few trades. Furthermore, as experienced traders continue to trade, they are also more likely to give feedback in general. Sellers are more likely to both reciprocate and give feedback unconditionally, and buyers are more likely to give feedback unconditionally. Overall, this is good news for eBay (and other trading systems that provide inter-partner performance feedback). As traders participate more in the system, they are more likely to be good citizens and hence provide feedback regardless of their trading partner's feedback giving actions.

Given our finding that new traders evolve into good citizens in terms of feedback giving, we expected there to be some "indoctrination" going on among eBay traders, but this does not appear to be true. Experienced sellers do not educate new buyers by giving them more feedback; neither do experienced buyers attempt to educate new sellers by giving them more feedback.

We also found that with high valued items, buyers are more likely to give feedback unconditionally but sellers are more likely to reciprocate. We speculate that as buyers care about the quality of high valued items more, they are more likely to pay attention and provide feedback. Experienced sellers would anticipate such behavior and strategically choose to reciprocate.

We also make a methodological contribution by building an econometric model to estimate feedback provision strategies in systems in which participants engage in bilateral interactions. Such types of systems can be electronic marketplaces, or systems that facilitate peer-to-peer sharing of services or resources. A special feature of our model is that the two parties' feedback provision strategies can be contingent on each other's actions, or not. Thus, either party can decide whether to give feedback based on what the other party does. With multinomial regressions, our model can be used to predict the participants' strategy choices based on their observable characteristics or the context of the interactions.

Chapter 3

Aggregation and Manipulation in Prediction Markets: Effects of Trading Mechanism and Information Distribution

3.1 Introduction

Market prices facilitate efficient resource allocations, and also act as information aggregators: they reflect market participants' valuation of resources (Hayek, 1945). Prediction markets — markets in which traders buy and sell bets on future events — are designed to explicitly take advantage of the information aggregation function of market prices to provide decision makers with forecasts of future events. Such markets have been created for a wide range of applications; examples include the Iowa Electronic Market for forecasting elections and other political events, the Hollywood Stock Exchange for forecasting movie box office receipts, and intra-company markets to forecast sales. In this paper, we use laboratory experiments to study the effectiveness of information aggregation of different variants of market scoring rule prediction markets, under differing information conditions. Our experimental results shed new light on the validity of theoretical predictions for these markets, as well as on the impact of common mechanism variations.

Although all prediction markets involve speculative bets on future events, the particular form taken by these bets can vary significantly. One common form is the *continuous double auction*, in which traders submit buy or sell orders for units of a security, and the market operator matches buy and sell orders to execute trades. When the outcome of the future event is known, the security is cashed out at a value that depends on the outcome. Continuous double auctions are very complex strategically, for the traders as well as the analyst. Recently, a new market form for prediction markets, the market scoring rule (Hanson, 2003), has become popular. Market scoring rules (MSR) are being used in a growing number of deployed prediction markets, including the public prediction market site *Inkling Markets* (inklingmarkets.com) that hosts public markets as well as closed markets for organizations, Yahoo!'s Predictalot[1], and Microsoft's internal prediction market (Berg and Proebsting, 2009). MSR markets have advantages over continuous double auction markets, particularly in situations with thin trade. In addition, they are more amenable to theoretical analysis, and there have been a number of recent studies that provide insight into optimal theoretical strategies in MSR markets (Hanson, 2003; Chen et al., 2007; Dimitrov and Sami, 2008; Chen et al., 2009). In this paper, we use human-subject laboratory experiments to study the speed and efficiency of information aggregation in MSR markets, while varying the mechanism form, constraints on trade timing, and information distribution pattern.

The first dimension of variation we consider is in comparing two commonly used mechanism forms that implement MSR markets: a direct mechanism in which traders report their beliefs as probabilities, and an indirect mechanism in which traders reveal their beliefs through buying and selling securities. The simplest representation of a market scoring rule market is as a sequence of *reported probabilities.* Each trade in the market involves a trader changing the current report. We call this type of market a *direct MSR* market. Once the outcome of the event is revealed, each trader is paid off according to a prespecified scoring rule, which depends on his report as well as the previous report. In practice, however, markets that use the market scoring rule, such as the public prediction market site Inklingmarkets.com, typically use an alternative mechanism: Traders buy and sell units of a security, but instead of trading directly with each other, they trade with an *automated market maker* who constantly adjusts the prices. We call this type of market an *indirect MSR* market. Direct and indirect MSR markets are formally equivalent, but may appear very different to traders in the market. There is an active debate about which interface is more effective in practice (Pennock, 2006); our laboratory experiments provide insight into this question.

Market speed and efficiency depends on appropriate behavior by traders, and hence trader strategies have to be taken into account. Hanson showed that MSR markets have a myopic honesty property: A trader trading only once maximizes her expected profit by reporting her true beliefs (Hanson, 2003). However, for traders

[1]See `http://labs.yahoo.com/project/336`, retrieved on July 14, 2010.

potentially using non-myopic strategies over multiple trades, the theoretical results show a sharp distinction based on the pattern of information distribution among the traders: If traders' information signals are independent *conditional on the true outcome of the event*, then the signals are substitutes, and honest reporting of beliefs is the optimal strategy even in a non-myopic sense; if the traders' information signals are *unconditionally independent*, then the signals are complements, and honest reporting is in general not a sequential equilibrium (Chen et al., 2009). In the latter case, a complete characterization of equilibria is unknown, but it is known that a trader can profitably deviate from the honest strategy profile by bluffing (trading in the opposite direction to her signal with some probability) or delaying (waiting for the other traders to reveal their information before trading); these deviations are construed as manipulative strategies. This motivates the second dimension along which we vary our experimental design: We study market performance under a complementary signal structure, and under a substitute signal structure. This enables us to paint a broader picture of the comparison between different market forms, as well as to conduct the first experimental test of these theoretical results on strategic manipulation.

The third variation we study is in providing the structure of strictly sequenced opportunities to trade, as compared to the standard approach of letting traders choose when to trade in an unstructured way. We have two motivations in considering a structure: First, the existing theoretical results (Chen et al., 2007; Dimitrov and Sami, 2008; Chen et al., 2009; Dimitrov and Sami, 2010) implicitly assume a structured order of trading opportunities, and this experiment allows us to test if this assumption is of practical significance. Second, enforcing more structured interaction has been shown to help in group forecasting performance (Graefe and Armstrong, 2008). Our experiments allow us to test if the additional structure of a trading sequence, which might simplify traders' information processing, improves the aggregation performance of a prediction market.

We designed and carried out market trading experiments to investigate the effect of varying these three dimensions on trader behavior and overall market performance. Our experiments involved 8 treatments generated by a factorial exploration of these three dimensions of variation. All experiments were conducted using markets with two traders in each market. Restricting the participation to two traders makes the signal interpretation problem for the subjects easier: there is only one information signal a trader does not have, and she can attribute every trade other than her own to the the other trader who has this information. It also allows for the market form to most closely match the theoretical models, thereby giving us a best-case situation in which to test the theoretical predictions. Based on the theoretical results summarized in Chen et al. (2009), we expect the following: In the substitutes markets, traders should trade honestly, as early as they can. In the complements markets, traders have an incentive to reveal their information as late as possible; they also have an incentive to bluff with some probability if they can correct the market later. Both these should lead to poorer early information aggregation in the complements case than in the substitutes case. Further, with ideal rational traders, the choice of direct or indirect mechanism should make no difference to the market aggregation. The comparison of structured and unstructured trading orders is an open question.

The results of our experiments make several contributions to our understanding of the aggregative and strategic properties of prediction markets with different trading mechanisms, and under different information conditions. First, we find that structured markets (with an exogenous sequence of trading opportunities), aggregate information more efficiently than unstructured markets, with an endogenous trading order. (This result was significant in three out of the four treatments. For the fourth treatment, the comparison was in the same direction, but not statistically significant).

Second, in the first experimental comparison between the direct and indirect trading mechanisms that have been proposed for market scoring rules, we find no significant difference in performance. Third, in testing the theoretical results on the effect of information distribution on manipulative strategies, we find that they are borne out for the structured market with exogenous sequence of trading opportunities, but not for the unstructured markets. This suggests that the timing and ordering of trades is an important feature to include in future theoretical research in this area.

The rest of this paper is structured as follows: In section 3.2, we summarize the prior research related to our work. Section 3.3 details our experimental design, analysis metrics, and hypotheses. We present the results in section 3.4. We summarize the paper and outline important directions for future work in section 3.5.

3.2 Related Work and Background

The theoretical underpinnings of using market prices as reliable forecasts of future events are provided by the theory of Rational Expectations Equilibrium (Muth, 1961; Radner, 1979; Fama, 1970). Rational expectations equilibrium models predict that, generically, prices in prediction markets can fully aggregate all individual traders' private information. Prediction markets' advantages over other methods of information aggregation such as polls and expert deliberations have also been empirically demonstrated in a large number of markets (Berg et al., 2008; Cowgill et al., 2008; Wolfers and Zitzewitz, 2004; Forsythe et al., 1992). Because of their perceived accuracy, as well as the fact that they are relatively easy and inexpensive to run, we are witnessing a rapid growth in the use of prediction markets as tools for information aggregation (Cowgill et al., 2008, Footnote 2).

3.2.1 Market Scoring Rules

In our study, we focus on market scoring rules (MSR) based prediction markets as suggested by Hanson (2003). Hanson outlined two alternative implementations of the MSR. One is a direct implementation of the MSR (we call it direct MSR), in which each trader reports their own predictions and receives payments accordingly. The other one is an indirect implementation of the MSR (we call it indirect MSR), which contains a market maker offering n securities each of which pays \$1 if the associated outcome is realized (Hanson, 2003; Chen and Pennock, 2007) and \$0 otherwise. The two implementations are mathematically and hence strategically equivalent, but they have very different look and feel to the market traders. Although both implementations have been used in practice (Pennock, 2006), to the best of our knowledge, there have been no empirical tests comparing the performance of these two implementations to provide guidelines for prediction market designers. It is one of our goals in this paper to compare the performance of these two implementations.

Direct MSR Scoring rules are tools for eliciting private beliefs. Given a random variable X which has n possible outcomes, to elicit an individual's, say Alice's, belief about the probabilities of each of these outcomes $p = (p_1, \cdots, p_n)$, we can ask her to express her beliefs by $r = (r_1, \cdots, r_n)$ — a vector of reported probabilities for the random variable X — and pay her based on the scoring rule $S = \{s_1(r), \cdots, s_n(r)\}$. Thus, if outcome 1 is realized, she will be paid $s_1(r)$; if 2 is realized, she will be paid $s_2(r)$, and so on. Alice maximizes her expected score $S(r)$ by choosing an r to report:

$$S(r) = \Sigma_{i=1}^{n} s_i(r) p_i \tag{3.1}$$

If the scoring rule is *proper*, Alice would find that $r = p$ maximizes her expected payoffs expressed in Equation (3.1). Popularly used proper scoring rules include quadratic, spherical, and logarithmic scoring rules.[2] In an MSR-based prediction market, traders report their forecasts sequentially, and have access to the sequence of forecasts made up to the current time. A trader earns the difference between her score and the previous trader's score. That is, if outcome i is realized, trader m who reported r_m will receive payment $s_i(r_m) - s_i(r_{m-1})$, where r_{m-1} is the report of the previous trader. Throughout this paper, we use an MSR based on the **logarithmic** market scoring rule.

Indirect MSR The market scoring rule can be viewed as a specific form of *automated market-maker*, an agent that posts prices, is always willing to trade securities at the posted price, and updates the prices following every trade. Thus, every trade in an MSR market is made with the automated market-maker as either the buyer or the seller; this is formally equivalent to the sequence model of direct MSR, while providing users with a mechanism that is more familiar to them from other markets. In particular, the (logarithmic) MSR equivalent market price of security i, p_i, can be derived from the following expression (Berg and Proebsting, 2009):

$$p_i = \frac{e^{s_i/b}}{\sum_k e^{s_k/b}} \tag{3.2}$$

where b is the scaling factor in the scoring rule, and s_i is the total amount of security i that has been sold. Berg and Proebsting (2009) detail the implementation of indirect MSR markets. (For reviewers' reference, the direct and indirect MSR mechanisms used in our experiment have been included in Figure B.1 and Figure B.2 in Section B.1 in the Appendix.)

3.2.2 Theoretical Analysis of MSR

A number of theoretical results have been shown concerning optimal strategies in market scoring rule markets. The first result, shown by Hanson (2003), is a myopic honesty result: A risk-neutral trader who trades only once (or does not consider any future trades while making a report) will maximize her expected utility by reporting her true belief about the item.

The strategic situation is more complex when traders can trade repeatedly, and are non-myopic. Two specific kinds of non-myopic strategies that have been analyzed (Chen et al., 2009) are dubbed as *bluffing* and *delaying*. In a bluffing strategy, a

[2]See Selten (1998) and (Cooke, 1991, p.139) for a discussion of various proper scoring rules.

trader first makes a trade that, with some probability, suggests information opposite to her true belief, so as to mislead other traders into reporting erroneous probabilities. In her next trade, she can then gain a profit by correcting the market price according to her true belief. Her total payoff will be the sum of the payoffs she earned from both trades. If she earns more from her second trade than she loses from her first trade, she gains a net profit. The extended game view of trade in a market also permits the delaying strategy: A trader with private information may choose to wait for other traders to report before revealing her private information. In comparison to the myopically optimal strategy, both bluffing and delaying have a negative effect on the speed of market convergence: The market price may not reflect the available information because one or more traders has either chosen to delay until later, or entered a report that is misleading.

In MSR-based markets, the profitability of these non-myopic strategies depends on the structure of traders' private information, i.e., the joint distribution of the signals they receive and the true outcome of the event. In particular, two natural distribution families have been studied: substitute and complementary signals.

In an information environment with substitute signals, private signals are independently distributed, *conditional on the true outcome.* For example, two people, A and B, try to predict if it is going to rain tomorrow. A tries to see if swallows fly low, and B uses the heuristic that "ring around the moon, rain is coming soon." Both A and B would receive private signals from "independent" sources about the weather tomorrow, though their signals are independent *conditional* on the current humidity of the air. In this case, the two signals are substitutes.

Another class of distributions involves signals that are (unconditionally) independent of each other. For example, firm A announces that at the end of the year each employee will receive a bonus if firm A's sales on both the East and West Coasts have met their targets. Employee E knows how the firm performed on the East Coast and employee W knows how it did on the West Coast. Assuming that the sales on the East and West Coasts are completely independent — knowing E's signals does not help one in guessing what W's signal is and vice versa. In this case, it can be shown that the signals are complementary: the predictive power of both traders' signals combined is greater than the sum of their individual predictive value.

When traders' private signals are substitutes, Chen *et al.* (Chen et al., 2009, 2007) show that misleading non-myopic strategies are not profitable. Honest reporting of beliefs at the earliest opportunity is the only perfect Bayesian equilibrium. On the other hand, when traders' private signals are complements, Dimitrov and Sami (Dimitrov and Sami, 2008; Chen et al., 2009) show that non-myopic players can indeed profit from deviating from honest reporting by either bluffing or delaying.

To the best of our knowledge, there has been no experimental study on non-myopic strategic manipulation in MSR-based prediction markets. Experimental tests of these predictions can not only provide guidance to the designers of prediction markets, but also inform theory development in terms of suggesting future directions. Our study is the first one to test the theoretical predictions of this literature.

3.2.3 Prior Experimental Work

There have been a number of studies conducted to measure the aggregative efficiency of prediction markets (see, for example, Plott and Sunder (1982) and Plott and Sunder (1988)). Apart from a few papers mentioned below, these have studied continuous double-auction markets or parimutuel markets, and not market scoring rule markets. We refer readers to the excellent literature review by Tziralis and Tatsiopoulos (2007) for further information.

There have been experimental studies comparing the accuracy of the forecasts produced by prediction markets and other information aggregation methods. Ledyard et al. (2009) study alternative forecasting techniques for combinatorial forecasting problems, and find that market scoring rules outperform all the alternatives studied. Graefe and Armstrong (2008) found that structured information aggregation methods, including prediction markets, perform better than the unstructured information aggregation method, i.e. face-to-face meetings. Healy et al. (2009) found that the relative performance of prediction markets to other alternative information aggregation methods, such as iterative polls, depends on the complexity of the information environment.

Our work is also related to prior work on manipulation in prediction markets. Hanson et al. (2006) gave half of the subjects (the manipulators) incentives to manipulate the market price by pushing the price up. Other traders were informed of the existence of these manipulators and the direction in which they wanted to push price. Hanson *et al.* found the market price was robust to manipulations, because knowing what the manipulators were trying to do, other traders effectively counteracted their influences. Oprea et al. (2007) further tested the influence of manipulators under the condition that all other traders only knew the existence of the manipulators but not the direction in which they push the price. They found that the traders still were able to counter-balance the manipulators' influence.

Our experiment differs from Hanson et al. (2006) and Oprea et al. (2007) in two aspects. First, we study internal manipulation — manipulations aimed at profiting within the same market — while Hanson et al. (2006) and Oprea et al. (2007) study external manipulation — manipulations aimed at profiting outside the market. To model the external manipulation, Hanson et al. (2006) and Oprea et al. (2007) gave the manipulators extra payments based on how successfully they influenced the market prices, in addition to their earnings as regular traders in the market. In our experiment, all the traders' payments are made as regular traders in the markets, even if they attempt to manipulate the market price. Second, our experimental prediction markets are based on a logarithmic market scoring rule, while Hanson et al. (2006) and Oprea et al. (2007) are based on double-auction markets. In our experimental design, the information distribution in the substitute treatment is consistent with the base model in Oprea et al. (2007), restricted to two traders.

3.3 Experiment Design

Our experiment follows a between-subject design — each subject only participates in one treatment. We recruited 256 subjects who were all students at the University of Michigan. Before the experiment began, an experimenter read the instructions to all the subjects. [3] These instructions included a tutorial on the experimental market's software interface, each individual's payoff functions, and the information they would receive based on the treatment. The experimenter then administered a paper-based quiz to all the subjects and checked each subject's answers in person. No practice rounds were given before the data collection started. Communications among the subjects were strictly forbidden. After the experiment ended, each subject filled out a short post-experimental survey about the strategies they used. [4]

In all the treatments, subjects participate in the market via computer software. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). Each trader started with 200 units of experiment currency in each round. They were also informed that 133 units of experimental currency could be later exchanged for U.S.$1.[5] For each treatment, we ran 4 independent sessions to achieve sufficient repetitions. There were 8 subjects in each session, which consisted

[3] The instructions can be found in the online supplement.

[4] The survey questions can be found in the online supplement.

[5] The average payment made to our subjects was $42, with the minimum being $28 and the maximum being $52.

	Direct MSR		Indirect MSR	
	Substitutes	**Complements**	**Substitutes**	**Complements**
Str.	Str-Dir-Sub	Str-Dir-Comp	Str-Indi-Sub	Str-Indi-Comp
Unstr.	Unstr-Dir-Sub	Unstr-Dir-Comp	Unstr-Indi-Sub	Unstr-Indi-Comp

Table 3.1: Experimental Design

of 25 rounds. At the beginning of each round, the 8 subjects were randomly paired into 4 groups of 2 traders. There are two traders in each market. This is not a typical setup of a market: markets usually have more than two traders. Nonetheless, we chose to study two-trader markets due to their simplicity — it is a good starting point to observe basic market dynamics. Future work is needed to test our results in markets with a larger number of traders.

We use a 2 x 2 x 2 factorial design as shown in Table 3.1. The factors are: direct vs. indirect MSR, substitute vs. complementary private signals, and structured vs. unstructured trading order. The following sections contain details about these treatments.

3.3.1 Structured vs. Unstructured

Almost all the prediction markets used in practice have unstructured participation, in that people freely choose the timing and frequency of their trades. However, the theoretical analyses of strategic behaviors in prediction markets are largely based on the assumption of exogenous ordering: people are given opportunities to trade in an order predetermined by factors beyond their control (Chen et al., 2007; Dimitrov and Sami, 2008; Chen et al., 2009). In these models, they are not forced to trade at every opportunity they receive; however, any timing-game elements of the strategic interaction between traders is abstracted away. In particular, traders are modeled as knowing when they will receive future opportunities to trade, as well as knowing that other traders have had opportunities to trade between their trades. It is unclear to what extent people would behave differently in these two types of environments (exogenous vs. endogenous trading order). An empirical test that compares these two types of markets, while keeping all other factors constant, will test the validity of this assumption, and can help guide future theoretical development.

Apart from testing the theory, this comparison may also influence market design. In small scale prediction markets in practice, it is often possible to impose some structure on the trading order. In fact, the use of structured information aggregation tool is not an entirely new idea. For example, in the 1950s, the Delphi method was developed as a multiple-round survey to elicit expert forecasts (Woudenberg, 1991). In each round, all participants are asked to provide their own forecasts and possibly comments. After each round, the aggregated forecasts are shown to all the participants, before they are asked to provide their revised forecasts again based on the aggregated forecasts of the group. The final forecasts are based on the aggregated forecasts in the final round. Our study is aimed at shedding light on whether structured prediction markets have advantages as well.

In treatments with a structured trading order, for each round we randomly determine the trading order between the two traders. The two traders then take turns to report their predictions; they may, of course, leave the previous report unchanged if they wish. In total, each trader has three turns to report. When it is a trader's turn, she has 30 seconds to make a decision. In treatments with an unstructured trading order, all traders can choose when to trade during the two-minute window in which the market is open.

3.3.2 Direct vs. Indirect MSR

We implement our MSR market using the logarithmic scoring rule, i.e., $s_i(r) = log(r_i)$, and call such type of market a *logarithmic market scoring rule* (LMSR) market. All the subjects report their prediction of the probability of the *black* ball being drawn in percentages.[6] Individual m's payoff from her report r_m is $200 \times (log_{10}(r_m) - log_{10}(r_{m-1}))$, where r_{m-1} is individual $m-1$'s report. Note that we used a scaler 200 to adjust the extent to which a subject can influence the market price. The initial market prediction is set to 50 (%). All the transactions in a market are displayed in real time to both participants.

In an indirect MSR market, subjects trade securities, each of which is based on a possible outcome of the random event. There are two securities in the markets, black and white, each paying one unit of our experiment currency if the corresponding outcome is realized, and zero units of experiment currency otherwise. The underlying market scoring rule and its parameters are exactly the same as those used in the direct MSR markets. To simplify the interface of an indirect MSR market, we only support trades in multiples of 10 and 50 shares. The exact prices of the shares and the new price after the transaction are shown to the subjects in real time. Restricting the

[6]Theoretically the range of probability should be 0% to 100%. But as the logarithmic function is undefined at 0, we restricted the probability predictions as integers in the range of $[1, 99]$.

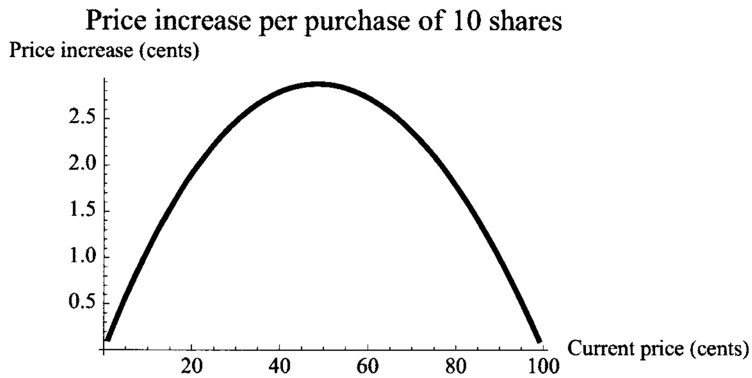

Figure 3.1: The price increase per purchase of 10 shares in an indirect MSR market.

number of shares per transaction in the indirect MSR markets might have an impact on the accuracy of the market predictions, but we argue that such an impact is likely small. Figure 3.1 illustrates how much increase in the market price can a purchase of 10 shares cause in an indirect MSR market. The highest price increase per 10-share purchase, 2.8 cents, occurs when the current price is 50 cents. As the current market price move away from 50 cents, the impact of a 10-share purchase decreases.

3.3.3 Substitutes vs. Complements

The information signals that traders see are generated according to a pre-specified distribution. We designed two information distribution treatments corresponding to the substitute and complement conditions as described in Section 3.2.2. In each treatment, all subjects try to predict the outcome of a hypothetical random event, based on their private information given by the experimenter, while sharing a common prior belief about the distribution of the outcomes.

For the substitutes treatment, we use the same substitutes signals as used in Oprea et al. (2007). At the beginning of each round, the computer randomly draws a black or white ball with equal probability. Once the round begins, each subject receives a private signal, either a "+" or a "-". The signal each subject receives depends on the color of the ball drawn at the beginning of the round. If the black ball was drawn, the signal will be a "+" with a 2/3 chance; and if the white ball was drawn, the signal will be a "-" with a 2/3 chance. At the end of each round, the color of the ball drawn is revealed, based on which all the subjects receive their individual payments.

In order to compare traders' behaviors in substitutes and complements markets, we specify an information condition with complementary signals such that the two are comparable. We impose the following two criteria on the complement signal structure to achieve a fair comparison:

C.1 The prior distributions of the random outcomes are the same under both environments. That is, the color of the ball drawn is black or white with equal probability.

C.2 The expected earnings of all the subjects are the same under both information conditions.

Criterion C. 1 ensures that the subjects in both treatments have the same prior belief about the security, to rule out confounds that people may behave differently when dealing with different prior probabilities. Criterion C. 2 implies that the value of the signals are equal. Specifically, if one person observes both subjects' private signals, in expectation, he would earn the same payoffs in both environments.[7]

We found the following complement signal structure that satisfies both criteria C. 1 and C. 2. Once the round begins, each subject receives a signal, a "+" or a "-" randomly drawn with equal probability. At the end of each round, a color (black or white) is randomly drawn by the computer, depending on the numbers of each signal that everyone together has received. How the signals determine the probability that a black ball will be drawn is shown in Table 3.2. For example, if there are 0 "+"'s and 2 "-"'s (there are only two traders in the market), there is a 19% chance that the black ball will be drawn, and an 81% chance that a white ball will be drawn. Note that, due to the different generating processes, the posterior probability of drawing a black ball given, say, two "+" signals, is slightly different in the complements treatment. It follows that, conditioned on getting two "+" signals, there is a different total expected profit from ideal aggregation in the substitutes and complements treatments. However, the probability of getting two "+" signals is also different, and by design, these two factors balance so that the expected total profit is the same in the two treatments.

3.3.4 Analysis Metrics and Hypotheses

We define *posterior efficient price* (PEP) as the prediction a perfect Bayesian who has observed all the signals in a round and has the correct prior belief would have

[7]Since we are using the logarithmic scoring rule, the fact that the two signals have the same value has an interpretation in information theory: in both settings, the two signals lead to the same reduction in the entropy (uncertainty) of the forecasted event.

# of "+"s	# of "-"'s	Prob Black	Prob White
0	2	19%	81%
1	1	50%	50%
2	0	81%	19%

Table 3.2: Mapping signals to the probability of the black or white ball being drawn

reached. In theory, a LMSR market with perfect information aggregation would converge to the PEP (Hanson, 2003). We will use the PEP as a benchmark to measure the accuracy of the forecasts made by our experimental prediction markets. We measure the market prediction accuracy as the mean squared error (MSE) of the market closing price to the PEP.

An alternative, and perhaps more natural, metric would have measured the distance of the final price from the actual realized outcome (i.e., whether the ball is black), or the correlation between the price and the realized outcome. In the long run, this metric would convey exactly the same information as the distance from the PEP. However, the actual outcome is subject to an additional layer of randomness, reflecting the distribution of the final outcome conditioned on the signal-pair of the traders. As we seek to measure the performance of the market in aggregating available information rather than serendipitously matching the true outcome, the PEP is a better comparison point: it yields the same long-run average performance, but for a finite number of rounds, the measure of the distance from the true outcome is noisier. In using the PEP, we are merely taking advantage of having controlled experiments rather than field trials.

With the analysis metric defined, we summarize our hypotheses and research questions below. First, we do not have any theory in predicting the effects of trading ordering on the performance of the prediction markets. Thus, we pose it as an open question:

- Do structured markets produce more accurate predictions than unstructured markets?

Second, since the direct and indirect MSRs are strategically equivalent (see Section 3.2.1), theoretically, varying the implementation of the MSR should not affect the accuracy of the market prediction.

Hypothesis 1. *The MSEs of the forecasts produced by direct and indirect MSR markets are the same.*

Third, theory predicts that there will be more delayed trading and bluffing in com-

	Direct MSR		Indirect MSR	
	Substitutes	**Complements**	**Substitutes**	**Complements**
Str.	265	289	267	345
Unstr.	413	385	446	386

Table 3.3: Market prediction accuracy comparisons

plements markets. Hence we have the following two hypotheses on traders' behavior:

Hypothesis 2. *There is more bluffing in complements markets than in substitutes markets.*

Hypothesis 3. *There are more delayed trades in complements markets than in substitutes markets.*

And at an aggregated level, behaviors predicted in Hypotheses 2 and 3 would lead to the following hypothesis:

Hypothesis 4. *Market prediction converges faster in substitutes markets than in complements markets.*

3.4 Experiment Results

3.4.1 Market Performance

Table 3.3 contains the MSEs of the market closing prices of all the treatments. Note that the lower the MSEs, the better the market prediction accuracy.

Structured vs. Unstructured Treating the four sessions of each treatment as independent data points, we conducted permutation tests to compare structured with unstructured markets while holding other conditions constant. Table 3.4 contains the results of the relevant tests. All the tests have the null hypotheses that the MSEs of the prices in both treatments are equal. For example, row 1 reads "under the condition of direct MSR and substitute private signals, the mean squared errors of the market closing prices in unstructured markets are higher than those in structured markets. The result is statistically significant at the 5% level (p-value = 0.028)."

We found that structured markets perform better than unstructured markets: under three out of the four conditions (rows highlighted in bold in Table 3.4) the

Condition	Alternative Hypotheses	P-value
Dir, Sub	**Unstr > Str**	**0.028**
Dir, Comp	**Unstr > Str**	**0.014**
Indi, Sub	**Unstr > Str**	**0.01**
Indi, Comp	Unstr > Str	0.27
Number of independent obs. per treatment		4

Table 3.4: Results of permutation tests comparing the MSEs of the market prices between structured and unstructured markets.

difference is statistically significant at the 5% level. Under the fourth condition, indirect MSR and complementary signals, although the difference is not statistically significant, it is in the expected direction: see column 4 in Table 3.3.

The differences between the MSEs of structured and unstructured markets are fairly large, as shown in Table 3.3. For example, in markets with direct MSR and substitutes signals, the MSE of the structured markets, 265, is significantly lower than that of the unstructured markets, 413 (see column 1 in Table 3.3). These results suggest that the additional structure provided by the fixed trading order improves traders' ability to interpret others' signals from their trades, and combine it with their own information. In an unstructured market, traders might not have the opportunity to fully consider others' signals before they trade. For example, both traders might submit their trades simultaneously. In a structured market, for each move, a trader has 30 seconds of dedicated time to learn about other traders' signals, based on which she then makes decisions on her own trades. These results are consistent with Graefe and Armstrong's finding that structured information aggregation methods work better than unstructured ones (Graefe and Armstrong, 2008).

Direct vs. Indirect MSR The test results related to comparing the two trading mechanisms are listed in Table 3.5, following the same format as in Table 3.4. Based on the results, we cannot rule out Hypothesis 1. We found no difference in the market prediction accuracy between the two mechanisms under all four possible combinations of the two information conditions and the two trading orders (all the tests in Table 3.5 are not statistically significant). As shown in Table 3.3, the actual MSEs

Condition	Alternative Hypotheses	P-value
Unstr, Sub	Direct $\neq$ Indirect	0.2
Unstr, Comp	Direct $\neq$ Indirect	0.8
Str, Sub	Direct $\neq$ Indirect	0.47
Str, Comp	Direct $\neq$ Indirect	0.21
Number of independent obs. per treatment: 4		

Table 3.5: Results of permutation tests comparing the MSEs of the market prices between markets with direct and indirect mechanisms.

of the market closing prices in these two mechanisms are very similar. In particular, in structured substitutes markets, the difference is 1 (265 vs. 267); in unstructured complements markets it is also 1 (385 vs. 386).

This result addresses an active debate in the prediction market research community about which mechanism is more effective. Within our controlled environment, there was no significant difference between the overall aggregative performance of the prediction market given the differing mechanisms. Further research needs to be carried out to determine if this is validated in field settings as well, as traders may not read and follow instructions as carefully in a field setting.

Substitutes vs. Complements In markets with a structured trading order, we found evidence supporting Hypothesis 4 — market prediction converges faster in substitutes markets than in complements markets. However this hypothesis is not supported in markets with an unstructured trading order.

Table 3.6 contains the results of the relevant tests. To establish a convergence result, we measured the MSE of the market price at two point in time: first right after both traders each had one chance to trade (marked as "early" in Table 3.6) and second when the market closes. We then compared the MSEs of the prices of a market at these two points across treatments. In structured direct MSR markets, we found no statistically significant difference between the MSEs of the market closing prices between the substitutes and complements markets (row 1 in Table 3.6, p-value = 0.23, two-sided test; and see row 3, p-value=0.16, two-sided test).

However when comparing the MSEs of the market prices right after both traders had had one turn to trade, we found the MSEs of the complements markets higher than those of the substitutes market (see row 2 in Table 3.6, p-value = 0.057, one-

Condition	Alternative Hypotheses	P-value
Str, Dir	Comp $\neq$ Sub	0.23
	Comp (early) > Sub (early)	0.057
Str, Indi	Comp $\neq$ Sub	0.16
	Comp (early) > Sub (early)	0.08
Unstr, Dir	Comp $\neq$ Sub	0.63
	Comp (early) $\neq$ Sub (early)	0.19
Unstr, Indi	Comp $\neq$ Sub	0.67
	Comp (early) $\neq$ Sub (early)	0.77
Number of independent obs. per treatment: 4		

Table 3.6: Results of permutation tests comparing the MSEs of the market prices between substitutes and complements markets.

sided test; and row 4, p-value = 0.08, one-sided test). We did not find a similar price convergence pattern in unstructured markets (see rows 5 - 8 in Table 3.6). One possible reason for this is that, in the unstructured markets, the increased difficulty of inferring others' signals (as evidenced by the worse overall aggregation performance) makes it more difficult to successfully execute a bluff-and-correct attack, or to exploit the greater value of one's information signal given the other trader's revealed information.

3.4.2 Strategic Behaviors

Bluffing

In structured direct MSR markets, we found support for Hypothesis 2 — there is more bluffing in complements markets than in substitutes markets. There were more groups whose first trades were inconsistent with their traders' private signals in the complements markets (23 out of 100 round/groups on average) than in the substitutes markets (15 out of 100 round/groups on average). Again we used a permutation test to compare the mean number of round/groups with dishonest first trades in both treatments, while treating each of the four sessions in each treatment as an independent data point. The test result shows that there was a statistically significant higher number of dishonest first trades in complements markets than in substitutes markets (p-value = 0.056, one-sided test). In all other markets (unstructured direct and indirect MSR, structured indirect MSR), however, the number of round/groups with dishonest first trades did not vary significantly between substitutes and complements markets.

Our survey data corroborates this finding. Table 3.7 shows the results of three linear regressions predicting the subjects' level of agreement to the statement: "To maximize my own profit, the best strategy is to report honestly according to my private information." The dependent variable takes five possible values in the set $\{-2, -1, 0, 1, 2\}$, with -2 indicating strongly disagree, and 2 indicating strongly agree. The independent variables include the dummy variables for the three treatment respectively. For example, *structured* takes a value of 1 if the subject was in a treatment with structured trading order. In regression (2) we also included three interaction terms among the three treatments. Regression (1) and (2) were run on a dataset containing data from all the sessions, and regression (3) was run on a subset of the data which includes only the treatments with structured trading order and direct MSR.

Our survey data suggest that in markets with complementary signals, traders are less likely to think honest trading is the optimal strategy. Regression (1) shows that across all the treatments, trading with others who have complementary signals leads to a 0.23 point (out of 4) reduction in the level of agreement to honest trading as the best strategy. In regression (2) we found that such a reduction is more pronounced in structured markets — the size of the reduction is 0.40, bigger than the main effect estimated in regression (1) (0.23). When we focus on the structured direct MSR markets in regression (3), we found an even bigger effect, a 0.56 point reduction in the subjects' level of agreement in honest trading being the best strategy.

Now that we have observed some amount of bluffing in both substitutes and complements markets, is bluffing profitable given the traders' reactions to each others' trades? In theory, we expect the incremental profit due to bluffing to be negative in the substitutes treatment. In the complements treatment, we expect the incremental profitability of bluffing to be 0 in equilibrium: if it were positive or negative, the bluffer could profitably increase or decrease the probability of bluffing accordingly. To answer this question, we conducted multivariate linear regressions on a panel dataset based on each subject's behavior in each round. To identify potential bluffing, for each group of each round, we only took the subject who trades first, to avoid the potential confound of the second trader trading dishonestly due to mis-interpreting

Dependent variable: level of agreement to honest trading as the best strategy			
	(1)	(2)	(3)
Structured	-0.062	0.125	
	(0.126)	(0.180)	
Complements	-0.234	0.094	-0.562
	(0.126)*	(0.222)	(0.237)**
Direct	-0.016	0.094	
	(0.126)	(0.226)	
Str × Comp		-0.406	
		(0.237)*	
Str × Direct		0.031	
		(0.237)	
Comp × Direct		-0.250	
		(0.237)	
Constant	0.250	0.094	0.344
	(0.133)*	(0.156)	(0.129)**
Observations	256	256	64
R-squared	0.014	0.028	0.077

Table 3.7: Linear regressions predicting the level agreement to honest trading as the best strategy
Notes: (i) Each column corresponds to a different regression, as detailed in the text. (ii) Standard errors are clustered at the session level. (iii) Standard errors in parentheses. * significant at 10%; ** significant at 5%.

previous traders' actions. Since for each treatment there are 4 groups in each round, 25 rounds in each session, and 4 sessions in total, we have a sample size of 400.[8]

Our dependent variable is the expected profit of the subject for the round given the trades she made, weighed by the posterior distribution of the outcomes. For example, suppose in a round with substitute signals, traders A and B both receive a "+" signal. Subject A makes a trade that changes the current market prediction from 50% to 65%. Her payoff will be $200 \times (log_{10}(0.65) - log_{10}(0.5)) = 23$ if the outcome is black, and $200 \times (log_{10}(0.35) - log_{10}(0.5)) = -31$ if the outcome is white. Given two "+" signals in the market, the posterior probability of the outcome being black is

$$\frac{\frac{2}{3} \times \frac{2}{3} \times \frac{1}{2}}{\frac{2}{3} \times \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2}} = 0.8 \tag{3.3}$$

and the posterior probability of the outcome being white is 0.2. Subject A's expected

[8]In the complements markets in some rounds neither of the traders traded, so there are only 388 data points.

Dependent variable: the subject's expected profit of the round				
	Sub(1)	Sub(2)	Comp(1)	Comp(2)
Gain	17.082	17.902	17.457	17.45
	(1.959)***	(1.822)***	(3.585)**	(3.606)**
Bluff	-2.349		-7.793	
	(-3.459)		(2.828)*	
BigBluff		-10.798		-8.101
		(2.237)**		(-3.644)
SmallBluff		10.916		-7.206
		(3.362)**		-4.27
Constant	192.441	192.011	192.615	192.61
	(0.479)***	(0.492)***	(2.008)***	(2.006)***
Observations	400	400	388	388
Number of Subjects	32	32	32	32
R-squared	0.126	0.152	0.149	0.149

Table 3.8: Linear regression predicting expected profit
Notes: 1) Fixed effect model and standard errors are clustered at the session level.
2) Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

profit from this trade then is $0.8 \times 23 + 0.2 \times (-31) = 12.24$. If subject A makes multiple trades in this round, her expected profit in this round would be the sum of her expected profit of all the trades she made in this round.

Our independent variables are the following:

- Gain — Dummy variable. It equals 1 if there is a profit to be gained in the market. Gain = 1 if $PEP \neq 50$ and Gain = 0 if $PEP = 50$.
- Bluff — Dummy variable. It equals 1 if the trade is inconsistent with the subject's private signal.
- BigBluff — Dummy variable. It equals 1 if the trade is a bluff and if it moves the current market price by more than 10 points.
- SmallBluff — Dummy variable. It equals 1 if the trade is a bluff and if it moves the current market price by less than 10 points.

We used a fixed effect model to account for repeated observations of each subject. The standard errors are clustered at the session level to adjust for intra-session influences among the subjects. Using fixed effects allows us to control for individual effects, which might be correlated with some of our independent variables, e.g., Bluff, BigBluff, and SmallBluff. The results are summarized in Table 3.8. For each market, we constructed models (1) and (2) to explore the main effects of bluffing and the effects of bluffing by different amounts, i.e., a small bluff (represented by variable SmallBluff) and a big bluff (represented by variable BigBluff).

In the substitutes markets, as an aggregate effect, we did not find a statistically significant effect of bluffing on expected profit of the round — we cannot rule out the hypothesis that the coefficient on Bluff, -2.349, is different from 0 by chance. However this does not mean that bluffing is not profitable in substitutes markets. When we break the Bluff variable down into BigBluff and SmallBluff, we found significant effects of both. If one bluffs by a large amount, her profit decreases by 10.8 points on average. But if she chooses to bluff by a small amount, she can earn about 10.92 points. Both coefficients are statistically significant at the 5% level. This result contradicts what theory predicts — if all players are perfectly rational and risk neutral, it is not profitable to bluff in substitute signal environments. However, in practice this is not true. Our subjects discovered that bluffing by small amounts is profitable. A possible explanation is that when bluffing, perhaps one does not have to move the price by very much to convince her opponent that she has a signal that is opposite to her true signal.

Contrary to expectations, our result from model Comp(1) shows that, on average a bluff is associated with a 7.8 point *loss* (statistically significant at 10% level). There are multiple potential explanations to this outcome. It could be that bluffers are bluffing too often, such that other traders stopped believing the information revealed. Hence bluffers are less likely to successfully execute their strategy. Or, it could be that traders are over-confident of their own private signals, such that they do not fully incorporate other traders' signals when they trade. In this case, bluffers will make a smaller profit than would be expected in theory. It could also be that, as a complex strategy, bluffing is more sensitive to errors in execution. Repeating the same analysis with big and small amount of bluffing, we did not find significant results in model Comp(2).

Delayed Trading

Again in the structured direct MSR markets, we found evidence supporting Hypothesis 3 — there are more delayed trades in complements markets than in substitutes markets. We restrict our analysis to the delaying behavior of the second mover — the trader who is randomly selected by the computer to move second in each round. This is because in our experimental setting, it is unclear what the first mover would gain by delaying her trades. There are six possible moves in each round. The first mover has the opportunity to trade during moves 1, 3, and 5. The second mover has the opportunity to trade during moves 2, 4, and 6. Since the last trading opportunity

Dependent variable: the actual first move of the second mover	
Complements	0.217
	(0.122)*
Constant	2.327
	(0.117)***
Observations	734
Number of Subjects	64

Table 3.9: Linear regression predicting the actual first move of the second mover
Notes: 1) We use a random effect model and standard errors are clustered at the session level. 2) Standard errors in parentheses. * significant at 10%; *** significant at 1%

is given to the second mover, if the first mover is to reveal her signal, she has to do so by the end of move 5.

For the second mover, we define a variable called ActualFirstMove, which is the move during which the second mover makes her first trade. The mean ActualFirstMove in the complements markets is 2.51, higher than the mean in the substitutes markets, i.e., 2.3. Treating each session as an independent data point, we conducted a permutation test to compare the mean ActualFirstMove between sessions of the complements treatment and sessions of the substitutes treatment. The result is marginally significant: p-value = 0.099, one-sided test. To test the robustness of this finding, we conducted a linear random effect regression on a dataset consisting of all second mover's first trades in each group of each round. The dependent variable is still the ActualFirstMove, and the independent variable, Complements, is a dummy variable which takes a value of 1 if the treatment is complements and 0 otherwise. A random effect model is appropriate here because the independent variable is exogenous — it is a treatment applied to randomized subjects. The result is shown in Table 3.9.

The coefficient on the complements variable is 0.217 (p-value = 0.07), and it is statistically significant at the 10% level. This result is consistent with what theory predicts — there are slightly more delayed trading in complements markets than in substitutes markets.

3.5 Conclusion

We conducted human-subject experiments to analyze the performance of variants of market scoring rule-based prediction markets, under differing information conditions.

We found that markets with structured trading orders provide better predictions than those with unstructured trading orders. This suggests a possible modification of prediction markets that may be beneficial when feasible, with small groups of traders. We also compared the performance of the direct and indirect MSR markets. We found no significant difference between the two forms.

Our experiments also enabled us to test theoretical predictions of strategic behavior in prediction markets. We compared the performance of markets under two different signal distribution conditions: when traders' signals are substitutes and when they are complements. In markets with a structured trading order, we found evidence supporting the theoretical prediction that there will be more manipulative behaviors, i.e., delaying and bluffing, in the complement signal markets than in the substitute signal markets. This was confirmed in our subject surveys, indicating that the subject group could perceive the strategic difference between the two environments. Interestingly, the theoretical predictions were not borne out in treatments without a structured order of trading opportunities. This result suggests that future theoretical work on the strategic behavior of prediction market traders will need to take into account the endogenous trading order typical of deployed prediction markets.

Our results suggest several important directions for future research. Firstly, our experiments were conducted in two-trader markets so that we could closely test the theoretical predictions and make cleaner inferences about trader behavior; it is important to confirm the validity of our conclusions with larger group experiments. Secondly, field experiments will complement our lab experiments comparing direct and indirect MSR: the effect of variations in interface may be more pronounced when users are not provided training in a controlled laboratory environment.

Chapter 4

Non-Monetary Mechanisms for the Provision of Excludable Public Goods

4.1 Introduction

Many information systems aggregate content contributed by their members and make it accessible to all members of the organization or community. The content is a public good. Public goods in general have the problem of under-provision (Samuelson, 1954a). Many solutions to this problem have been proposed by economists (Groves and Ledyard, 1977; Walker, 1981; Bagnoli and Lipman, 1989; Admati and Perry, 1991). These solutions all require the use of monetary payments among the individual participants. Such solutions are not well suited to many online information systems, for which monetary payments are not practical for either economic or social reasons.

Information technology makes exclusion a potential instrument for eliciting content contribution without the use of money. For networked computational information systems, access control is relatively cheap, with varying degrees of identification possible using passwords, Internet Protocol (IP) addresses or cookies. Contributions can be automatically monitored (at least by contributor and by file size), creating the possibility of algorithmic exclusion based on contribution. Contribution quality can be evaluated and quantified either by human raters, or algorithmically in some cases. Examples of human ratings include user ratings of book reviews on Amazon, of answers on Yahoo!Answers, and of comments posted on Slashdot.org; Lampe and Resnick (2004) found that the Slashdot peer-rating system works fairly well. As an algorithmic example, the value of each edit on Wikipedia can now be measured by how long it survives subsequent deletions or modifications (Adler and de Alfaro, 2007).

In fact, exclusion-based mechanisms have been used in practice. The popular Bittorrent P2P protocol implements an exclusion-based system. One can only download file if she uploads as well (Cohen, 2003). Some peer-to-peer (P2P) file sharing communities, e.g., *ilovetorrents.com*, use a simple exclusion rule: it requires its members to maintain a minimum upload/download ratio to be eligible for continuing to download files.[1]

These mechanisms do not require the use of monetary transfers. They are simple to implement and seem to perform reasonably well in existing file-sharing systems. However, most of these mechanisms used in practice are implemented in an *ad hoc* manner. There has been little systematic study on the strategic properties of these mechanisms. Usually the site administrator makes an arbitrary decision on the parameters of these mechanisms and the rest of the participants would have to follow the rules. How do we choose among various exclusion (or other) mechanisms to obtain the best social welfare, or the highest quality-adjusted contribution quantity? For a given mechanism, how do we set the parameters to achieve one of these goals?

As a first step towards understanding these simple exclusion-based mechanisms, I use game theory to analyze the performance of two mechanisms: the minimum threshold mechanism, under which one can only access the public goods if her contribution is higher than a pre-specified threshold, and the ratio mechanism, under which a user consumes at most an amount proportional to her own contribution level. I derive equilibrium predictions for these two mechanisms and analyze their performance in terms of social welfare. My results indicate some advantages of the minimum threshold mechanism over the ratio mechanism. There exist some conditions under which the minimum threshold mechanism can achieve the social optimum, but the ratio mechanism cannot. Furthermore, if the ratio mechanism implements a no-exclusion equilibrium, the same outcome can always be implemented by the minimum threshold mechanism.

4.2 Exclusion-based mechanisms

The majority of the exclusion-based mechanisms proposed in the literature were to solve the problem of cost sharing (Moulin, 1994; Deb and Razzolini, 1999; Young, 1998; Bag and Winter, 1999; Feldman et al., 2004). Moulin (1994) first analyzed the serial cost sharing (SCS) mechanism, under which all individuals who enjoy the same

[1]See `http://www.ilovetorrents.com/rules.php`, retrieved on Dec 11, 2008.

amount of goods share the cost of producing these goods equally. The SCS mechanism is strategy-proof, fair, and individually rational, but it does not guarantee efficient outcomes. Others have proposed mechanisms that can achieve the socially optimal outcomes using strong Nash implementation (Young, 1998) and subgame perfect Nash implementation (Bag and Winter, 1999).

The SCS mechanism has been put to test in human-subject laboratory experiments in the context of public goods provision (Gailmard and Palfrey, 2005).[2] Gailmard and Palfrey (2005) compare the performance of SCS with two other voluntary provision mechanisms that do not involve exclusion: voluntary cost sharing with proportional rebates (PCS) and without rebates (NR). The type of public good used in their experiment is a public project, e.g., a bridge or a park, which requires a fixed amount of funds to produce. If enough funds are collected the public good is produced, otherwise it is not. In case there are excess pledges, under PCS, rebates are paid to contributors based on the proportion of their individual contributions, while no refunds are provided under NR. Gailmard and Palfrey (2005) found that SCS is outperformed by PCS in terms of both consumer surplus and social efficiency. They attribute this difference to two main sources of inefficiency in SCS. For one thing, PCS does not suffer from *exclusionary inefficiency* — inefficiency created by simply excluding some participants while including them could have been costless — because it is not exclusion based, whereas under SCS exclusions do occur hence hurting the efficiency. Also, PCS allows unequal cost shares among individuals who consume the same amount of public goods, thus allowing high value participants to subsidize low value participants. Under SCS, subsidies among individuals are not possible.

The work closest to ours is Wash and MacKie-Mason (2009). While assuming that all the individuals can be ordered by their marginal net utility of the public goods, Wash and MacKie-Mason analyzed the minimum threshold mechanism and derived similar results to my Fact 3 — which characterizes a Nash equilibrium under the minimum threshold mechanism — and Fact 4 — which shows that the minimum threshold mechanism can improve the efficiency compared to a voluntary contribution mechanism. I show Fact 3 and 4 in an environment under which this order-able net marginal benefit assumption also holds, and then build on these two facts to further analyze the minimum threshold mechanism and compare it with the ratio mechanism.

[2]The SCS mechanism has also been tested in laboratories in the context of cost sharing for private goods (Chen, 2003; Razzolini et al., 2007).

4.3 Setup and Notation

Let us consider N individuals jointly producing some excludable public goods. For simplicity, I assume that it is impossible for each individual to produce the good standalone. Each individual has an initial endowment, ω ($\omega > 0$). I assume ω is sufficiently high to not be a binding constraint for our participants' utility maximization. An individual i who contributes x_i amount of public goods has the following utility:

$$u_i(x_i) = \alpha_i v(x_{-i} + x_i) + \omega - c(x_i) \tag{4.1}$$

- α_i is the "value coefficient" for individual i.
- v and c are continuous functions, with $v' > 0$, $v'' < 0$, $c' > 0$, and $c'' > 0$. $v(0) = c(0) = 0$ and $v'(0) > c'(0) > 0$.
- x_{-i} is the sum of everybody's contribution except i's.
- $F(\alpha)$ is the *realized* distribution of α on the support of $[\underline{\alpha}, \bar{\alpha}]$. $\underline{\alpha} > 0$.

The v and c functions are common knowledge among all the participants and the social planner. Everybody, including the social planner, also knows the distribution of the value coefficient (α) but does not know each individual's α_i. In this system, an outcome is specified by a vector of contribution $X = \{x_1, x_2, \cdots, x_N\}$, where x_i is participant i's amount of contribution, and a vector of consumption $Z = \{z_1, z_2, \cdots, z_N\}$, where z_i is participant i's amount of consumption.

Given that the cost function is convex, in the socially optimal outcome, each individual would contribute an equal amount towards to the optimal level of public goods, y^{FB}, and consume the full amount of y^{FB}. y^{FB} satisfies the following condition:

$$y^{FB} = \arg\max_{y} v(y) N \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha) \mathrm{d}\alpha - Nc(y/N) \tag{4.2}$$

The social optimum that results from the above welfare maximization might violate individual rationality: some individual with a low α might be receiving a negative payoff. Alternatively, one could derive different benchmarks for measuring efficiency by imposing individual rationality as a constraint.

I consider two mechanisms, the minimum threshold mechanism and the ratio mechanism, while comparing them against the voluntary contribution mechanism. These three mechanisms are formally defined as follows:

Voluntary Contribution Mechanism (VCM) Every individual i chooses x_i voluntarily, and consumes the full amount of the public good.

Minimum Threshold Mechanism (MTM) Let t be a publicly known threshold. If $x_i \geq t$, then agent i is included and can access the full amount of the public goods produced; otherwise i is excluded and earns a payoff of zero.

Ratio Mechanism (RM) Let r ($r > 1$) be a publicly known ratio. An individual who contributes x_i can access $min(rx_i, x_{-i} + x_i)$ amount of the public goods.

4.4 Voluntary Contribution Mechanism

In preparation for analyzing MTM and RM, I first characterize the VCM equilibrium as a benchmark. Fact 2 establishes the existence of a unique Nash equilibrium (NE) under VCM, and Fact 1 highlights that in equilibrium individuals' contribution level monotonically increases in their value coefficients, α.

Fact 1. *Under the voluntary contribution mechanism, in a Nash equilibrium individuals' contribution levels weakly increase in α.*

Proof. See appendix C.1. □

Fact 2. *There exists a unique pure strategy Nash equilibrium in the voluntary contribution mechanism.*

Proof. See appendix C.2. □

4.5 Minimum Threshold Mechanism

In this section, I first characterize an MTM's Nash equilibria. I also show that MTM can improve the social welfare compared to VCM, and under certain conditions it can even reach the social optimum. Last, I discuss MTM's potential of leading to over-production. Fact 3 and Fact 4 have been proven in a slightly different context in Wash and MacKie-Mason (2009) (see Proposition 1). Here I prove it in my own setup to facilitate further analysis of the MTM.

4.5.1 Equilibrium predictions

In this section, I show that in a Nash equilibrium, individuals are divided into two main categories: the non-contributors and the contributors. The non-contributors find it not worthwhile joining the "club", and they would rather contribute nothing and consume none of the public good. These non-contributors all have a value coefficient less than or equal to α_t. The individual with α_t is indifferent between contributing nothing and thus being excluded, and contributing exactly t and thus having full access to the public goods produced. All the participants who have a value coefficient greater than α_t would be willing to join the club by contributing at least t. There are two types of contributors in a typical equilibrium: those who contribute exactly t, and those who contribute more than t. For the contributors with their value coefficients less than α_m but greater than α_t, although it is not worthwhile for them to contribute more than t, they are willing to contribute the minimum amount that allows them the full access to the public goods. The contributors with their value coefficients greater than α_m, however, would not only want to be included in the club, but also find it in their own benefits to increase the total amount of production by contributing more than the threshold t. Again, the individual with α_m is indifferent between contributing exactly t amount or contributing infinitesimally more than t.

Fact 3. *In a Nash equilibrium, under the MTM mechanism with a threshold t ($t > 0$), individuals' contributions weakly increase in α. In an equilibrium with y^{MTM} being the total amount of public goods produced ($y^{MTM} > 0$), an arbitrary individual, i, contributes x_i:*

$$x_i = \begin{cases} 0 & \textit{if } \alpha_i \le \alpha_t \\ t & \textit{if } \alpha_t < \alpha_i \le \alpha_m \\ x_i^* & \textit{if } \alpha_i > \alpha_m \end{cases}$$

where $\alpha_t = c(t)/v(y^{MTM})$ and $\alpha_m = c'(t)/v'(y^{MTM})$.

Proof. See appendix C.3. □

Remark: Fact 3 describes a typical equilibrium, meaning an "interior" solution: when both α_t and α_m fall in the interval $[\underline{\alpha}, \bar{\alpha}]$. Some Nash equilibria under MTM might be corner solutions. For example, if $\alpha_m > \bar{\alpha}$, no one would contribute more than the threshold in equilibrium. By the same token, if $\alpha_t < \underline{\alpha}$, no one would be contributing zero in equilibrium, hence no one is excluded under the MTM. Exactly what kind of equilibrium will be realized depends on the individuals' utility functions, their cost functions, and the distribution of the value coefficient.

Remark: The equilibria characterized by Fact 3 need not to be unique. In principle, given a particular set of v and c functions, a distribution of α, and the total number of participants, one can solve for all the Nash equilibria by constructing equations of the following form:

$$N \int_{\underline{\alpha}}^{\bar{\alpha}} x_\alpha(y,t) f(\alpha) \mathrm{d}\alpha = y$$

where y is the equilibrium level of production, $x_\alpha(y,t)$ is the amount of contribution by the individual with α, as a function of y and t. One can derive the functional form of x_α based on the equilibrium characterization provided by Fact 3.

4.5.2 Efficiency

A first question about MTM is whether it can improve efficiency on VCM. Fact 4 establishes that if the VCM does not achieve the social optimum, we can always find an MTM with threshold t that increases the social welfare. Wash and MacKie-Mason prove a similar result while requiring there to be a large number of participants (see Proposition 2 in Wash and MacKie-Mason (2009)).

Fact 4. *If under VCM a non-zero but not efficient amount of public goods were provided, we can always find a t ($t > 0$) such that the corresponding MTM has an equilibrium that strictly increases public good production. Further, if the equilibrium level of public good production under MTM is not greater than the socially optimal amount, social welfare must have increased.*

Proof. See appendix C.4. □

Intuitively, the reason why sometimes MTM can be more efficient than VCM is simple: VCM can be seen as a special case of MTM with $t = 0$. So MTM should perform at least weakly better. Fact 4, shows further that as long as VCM does not reach the social optimum, MTM can always *strictly* increase social welfare compared to VCM.

If the social planner only wants to increase the production level of the public goods compared to VCM, then Corollary 1 shows that he only needs to keep increasing t until the lowest contributor (not the participant with the lowest value coefficient) under VCM is indifferent between contributing an amount t and being excluded.

Corollary 1. *For all α profiles, as long as the MTM equilibrium does not exclude any contributors in the VCM and it makes at least one agent increase contribution compared to VCM, the total amount of public goods produced must have increased compared to VCM.*

Proof. Let α_t be the value coefficient of an individual who contributes the least among all the contributors. Replacing $[\underline{\alpha}, \bar{\alpha}]$ with $[\alpha_t, \bar{\alpha}]$, the proof of Fact 4 establishes Corollary 1. $\square$

Now we know that MTM can improve the efficiency compared to VCM, but can it achieve the social optimum? The answer is "sometimes". Formally, Proposition 1 provides the necessary and sufficient condition for the existence of a socially optimal Nash equilibrium under MTM.

Proposition 1. *There exists a Nash equilibrium under an MTM with threshold $t = y^{FB}/N$ such that the socially optimal outcome is achieved, if and only if the following condition holds:*

$$\underline{\alpha} \geq \frac{c(\frac{1}{N}p^{-1}(N\int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha))}{v(p^{-1}(N\int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha))} \tag{4.3}$$

where $p(x) = \frac{c'(\frac{x}{N})}{v'(x)}$.

Proof. See appendix C.5. $\square$

Using equation (4.3), I characterize the relationship between $\underline{\alpha}$ and $N\int_{\underline{\alpha}}^{\bar{\alpha}} f(\alpha)\mathrm{d}\alpha$ when the socially optimal outcome is achievable. More specifically, I provide a bound for how low any participant's value coefficient can be, based on the sum of all the participants' value coefficients, such that the social optimum is still achievable.

How does this bound, i.e., the right hand side of equation (4.3), vary with the total number of participants? In Proposition 2, I show that, if the v function is sufficiently concave, and the realized distribution of α is stationary as the number of participants increases, this lower bound decreases in N. That is, if the marginal value of public goods decreases fast enough, increasing the total number of participants will decrease the lower bound for α. When this lower bound decreases, more potential participants, especially those with low α, can be attracted to both contribute to and consume the public good.

Proposition 2. *Assuming the realized distribution of α, $F(\alpha)$, is stationary as N increases, if the v function satisfies the condition that $v'(y) + yv''(y) < 0$ for all $y > 0$, then the right hand side of equation (4.3) decreases in N.*

Proof. See appendix C.6. $\square$

Intuitively, this comparative static reflects two effects of increasing the total number of participants. On one hand, if there are more people to share the cost of production, each individual needs to contribute less in order to access the public goods. I call this the "cost-sharing" effect. On the other hand, if there are more participants, the total value generated by each additional unit of public goods increases, leading to an increase in the socially optimal amount of public goods. All else being equal, increasing the socially optimal amount of public goods will lead to increased individual contribution (recall that in a socially optimal outcome the cost is shared equally among all the participants). I call this the "bigger pie" effect. Whether each individual's cost of contribution in equilibrium increases depends on which one of these two effects dominates.

If the value function, v, is sufficiently concave, then the "bigger pie" effect is dominated by the "cost-sharing" effect and thus it costs less for each individual to "stay in" as the total number of participants, N, increases. Given that the socially optimal amount of public goods increases in the number of participants, the individual with $\underline{\alpha}$ can contribute less and consume more. Thus, as the number of participants increases, as long as the distribution of α stays stationary, the system can potentially include people with even lower α.

If instead the v function is gently concave, then the "bigger pie" effect dominates the "cost-sharing" effect. Thus when the total number of participants increases, each individual's cost of being included also increases. In the mean time, the socially optimal amount of public goods would have also increased, thus increasing each individual's value of being included. Exactly whether the lower bound for α increases or decreases in this case depends on the relative magnitude of the increases in value and and the increases in the cost of staying in.

4.5.3 Over-production

In general, the VCM does not lead to over-production. The fact that participants are optimizing their own utility implies that the last unit of contribution they made still created more value than its cost.[3]

[3]Of course, I am not considering cases in which some contributors are motivated by reasons other than maximizing their direct utility from consuming the goods. For example, some people write Wikipedia articles to polish their writing skills. If we take into account how much fun a contributor experiences from, for example, writing the 1134th review for a movie on IMDB.com, then it is incorrect to call it "over-production".

The MTM, however, has the potential for over-production.[4] Under MTM, some individuals are motivated to contribute because they need to "pay" for accessing the public goods. Given that the marginal social value of public goods decreases, if in a system a large number of participants all pay the minimum amount in order to gain full access to the public goods, it can lead to over-production.

In Fact 1 I have characterized the conditions under which MTM can achieve the socially first-best outcome. If we replace the "$\geq$" in equation 4.3 with a strict "$>$", we obtain sufficient conditions under which over-production can occur under MTM. This condition is only sufficient because it is restricted to equilibria in which no exclusion occurs. Certainly, over-production can also occur when there is some degree of exclusion in equilibrium.

4.6 Ratio Mechanism

Before I derive the Nash equilibrium of a ratio mechanism (RM), I first define a term, the *r-optimal amount of contribution.* Given an RM with r, an individual with α_i maximizes her utility by contributing x_i^r, provided that the total amount contributed by all the participants is greater than rx_i^r. I call x_i^r individual i's r-optimal amount of contribution. More specifically, x_i^r satisfies the following condition:

$$\alpha_i r v'(r x_i^r) - c'(x_i^r) = 0 \tag{4.4}$$

In a Nash equilibrium, with an amount of public good production y^*, any participant with a value coefficient sufficiently low (lower than a threshold I call α_m) contributes her r-optimal amount. This is because the total amount of public good produced, y^*, is greater than r times her r-optimal amount of contribution. Thus she can take full advantage of the amplifying effect of r: compared to a situation in which an individual produces and consumes the public good standalone, the RM mechanism "amplifies" her consumption by a factor of r. For a participant with a higher value coefficient (higher than α_m), however, the total amount of public good produced cannot satisfy her demand — r times her r-optimal amount. In this case her level of consumption is constrained by y^*, and she could maximize her utility by

[4]Wash and MacKie-Mason (2009) also highlighted the possibility that MTM can lead to over-production, albeit for a different reason. Using the example of Wikipedia, they pointed out that some participants might be motivated to contribute a lot of articles for reasons that are not modeled in my paper, e.g., altruism or fun. In this case, MTM would not help in increasing the social welfare because the public goods might have already been over-produced.

contributing exactly y^*/r. She could also consider contributing more than y^*/r, if doing so increases her utility. Note that if she contributes x units more than y^*/r, the extra amount x cannot be amplified by r. Now her x units of extra contribution can only increase her consumption by x unites, not rx units. I found that if one's value coefficient crosses a threshold, α_k, she would indeed profit from contributing more than y^*/r. Otherwise, she maximizes her utility by contributing exactly y^*/r. I summarize this characterization of a NE under RM formally in Proposition 3.

Proposition 3. *For a ratio mechanism with any given r ($r > 1$), there exists a pure strategy Nash equilibrium under which the total amount of public goods produced is y^*, and an individual i with α_i will contribute x_i^*:*

$$x_i^* = \begin{cases} \arg\max_x \alpha_i v(rx) - c(x) & \text{if } \alpha_i \leq \alpha_m \\ y^*/r & \text{if } \alpha_m < \alpha_i \leq \alpha_k \\ \arg\max_x \alpha_i v(x_{-i}^* + x) - c(x) & \text{if } \alpha_i > \alpha_k \end{cases}$$

where $\alpha_m = \frac{c'(y^*/r)}{rv'(y^*)}$ *and* $\alpha_k = \frac{c'(y^*/r)}{v'(y^*)}$.

Proof. See appendix C.7. □

Remark: The equilibrium described in Proposition 3 may not be unique. I have not been able to find necessary and sufficient conditions under which the RM has a unique pure strategy Nash equilibrium. However, given a particular pair of v and c functions, a distribution of α, and the total number of participants, it is straightforward to solve for all the Nash equilibria by constructing equations of the following form:

$$N \int_{\underline{\alpha}}^{\bar{\alpha}} x_\alpha(y, r) f(\alpha) \mathrm{d}\alpha = y$$

where y is the equilibrium level of production, $x_\alpha(y, r)$ is the amount of contribution by the individual with α, as a function of y and r. One can derive the functional form of x_α based on the equilibrium characterization provided by Proposition 3.

I next explore the question: can the RM achieve the social optimum, at least sometimes? The answer is "no". I found the RM mechanism can never achieve the social optimum in the setting of this paper as detailed in section 4.3.

Proposition 4. *The ratio mechanism can never achieve the social optimum, as long as there are at least two participants with different value coefficients.*

Proof. If an RM achieves the socially first-best outcome, it produces y^{FB} amount of public goods with each individual contributing exactly y^{FB}/N, implying that r has to be set to equal N. As individuals' levels of contribution increases (weakly) in α, a necessary condition for all the individuals to be willing to contribute y^{FB}/N is that the person with the lowest α, $\underline{\alpha}$, finds it optimal to do so. Formally, it requires $\underline{\alpha}Nv'(y^{FB}) - c'(y^{FB}/N) \geq 0$. Since y^{FB} is the first best production amount, it satisfies the following first order condition: $\int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha N v'(y^{FB}) - c'(y^{FB}/N) = 0$, implying $\underline{\alpha}Nv'(y^{FB}) - c'(y^{FB}/N) < 0$ as long as $\underline{\alpha} < \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha$. □

The intuition behind Proposition 4 is as follows. At the social optimum, an extra unit of public goods produced creates some marginal value for all the participants that equals the marginal cost. The marginal value created for all N participants can also be seen as the marginal value created for N participants all with an average α. Suppose an RM equilibrium reaches the social optimum, which implies for the individual with $\underline{\alpha}$, any extra unit of contribution she makes creates some marginal value for *herself* that equals the marginal cost. Since it is under an RM with $r = N$, the marginal value of each unit of contribution for an individual is amplified by N. Thus for the individual with $\underline{\alpha}$, any extra unit of contribution she makes can be seen as creating some marginal value for N participants all with $\underline{\alpha}$. Obviously, any amount of production creates more value for N average participants than for N low-value participants. Thus for the individual with $\underline{\alpha}$, any unit of contribution would create less value than it costs. She therefore has no incentive to contribute up to y^{FB}/N under the RM.

Now, we know that RM cannot achieve the socially optimal outcome. It would still be useful to characterize how the social welfare varies in r. This is a non-trivial task. We saw in Proposition 3 that in equilibrium, individuals' contribution level relates to the total amount of production in more than one ways, which makes it hard to derive the social welfare as an explicit function of r. Further research is needed to either simplify the problem, or discover new ways of welfare maximization. Here, as a first step towards understanding the effect of varying r, I introduce a lemma that characterizes how an individual's r-optimal amount of contribution varies with r.

Lemma 1. *Under RM, individual i's r-optimal contribution x_i^r increases in r if and only if the v function satisfies the following condition:*

$$v'(x_i^r) + x_i^r v''(x_i^r) > 0. \tag{4.5}$$

Otherwise x_i^r decreases in r.

Proof. See appendix C.8. □

Remark: Equation (4.5) implies that the v function is "gently" concave around x_i^r. We know for the log function, $v(x) = log(x)$, $v'(x) + xv''(x) = 0$. That is, the condition as expressed by equation (4.5) is requiring the v function to be less concave than the *log* function around x_i^r.

Remark: Equation (4.5) only needs to hold in the neighborhood of x_i^r for individual i's r-optimal amount of contribution to increase in r. To generalize, a natural sufficient condition for the r-optimal amount of contribution to be increasing in r for all the participants is for the v function to be globally "gently" concave.

4.7 Welfare Comparisons

Although MTM and RM both are exclusion-based mechanisms, there may not always be exclusion in equilibrium under these mechanisms. Under both MTM and RM, there are equilibria in which every participant has access to the total amount of public goods produced. This can be verified by the fact that Fact 3 and Proposition 3 both can contain corner solutions for the equilibrium. Equilibrium outcomes that do not involve exclusion might be desirable when not resulting in exclusion in equilibrium is a design constraint. Here I provide a result that compares MTM and RM when they both lead to no-exclusion equilibria.

Proposition 5. *If there exists a Nash equilibrium under the RM such that everyone has access to the total amount of public goods produced, then there also exists a Nash equilibrium under the MTM that achieves the same outcome.*

Proof. See appendix C.9. □

Proposition 5 highlights another area in which MTM out-performs RM. The cases under which RM can achieve a no-exclusion equilibrium is a subset of the cases under which MTM can achieve the same.

Proposition 5, together with Proposition 1 and 4, does not complete the welfare comparison between MTM and RM. There are situations under which neither mechanism can implement a no-exclusion equilibria. Insights on how MTM compares with RM would still be useful for mechanism design. I will leave it as future work.

4.8 Conclusion and discussion

In this paper I examined two simple exclusion-based mechanisms, the minimum threshold mechanism and the ratio mechanism, in a complete information environment. I characterized the Nash equilibria for these two mechanisms, and analyzed their performance in terms of social welfare. The results I obtained so far indicate that the minimum threshold mechanism has some advantages over the ratio mechanism. There exists some conditions under which the minimum threshold mechanism can achieve the social optimum, but the ratio mechanism can never achieve the social optimum. Furthermore, if the ratio mechanism implements a no-exclusion equilibrium, the same outcome can also be implemented by the minimum threshold mechanism.

These results, together with Wash and MacKie-Mason (2009), are a first step towards understanding these simple exclusion-based mechanisms for the public good provision in large information systems. The results I derive are limited in a number of aspects. First, both MTM and RM might lead to multiple equilibria, which makes the welfare analyses of these mechanisms complex. Further studies are needed to characterize the efficiency of these mechanisms in the presence of multiple equilibria. Second, I have not completed the welfare comparisons between MTM and RM. Although I have shown that under certain conditions MTM Pareto dominates RM, there might be conditions under which RM can achieve higher social welfare than MTM. Future work is needed to characterize these conditions. Last, I studied these mechanisms in a complete information environment, which requires both the social planner and the participants to know the distribution of the participants' valuation of the public goods and their cost functions. Future research is needed to explore the performance of these mechanisms when the participants and/or the social planner might have limited information.

Chapter 5

Conclusion

I have reported three studies centering on motivating individuals to provide high quality information to online information aggregation systems, also known as social computing or Web 2.0 systems.

In my first study, my co-authors Jeffrey MacKie-Mason and Paul Resnick and I developed an econometric model to study the feedback provision strategies used by participants in systems for bilateral interactions between strangers. We then applied our model to analyze the feedback provision strategies of eBay traders. We hypothesized that three types of feedback provision strategies were played by the traders: give (feedback) unconditionally, abstain (from giving feedback) unconditionally, and reciprocate. We found that all three types of strategies were being played by the traders. In particular, we found that in a substantial faction of cases, traders were strategic feedback reciprocators — 23% of the buyers and 20% of the sellers. We argue that the knowledge about the existence of these reciprocators may be a motivation for some traders to give feedback unconditionally, as they anticipate their partners to reciprocate.

In addition to the empirical findings, we also make a methodological contribution by building a simultaneous equation maximum likelihood model to estimate latent feedback provision strategies in systems in which participants engage in bilateral interactions. Such types of systems can be electronic marketplaces, or systems that facilitate peer-to-peer sharing of services or resources. A special feature of our model is that the two parties' feedback provision strategies can be contingent on each other's actions, or not. Thus, either party can decide whether to give feedback based on what the other party does. With multinomial regressions, our model can be used to predict the participants' strategy choices based on their observable characteristics or the context of the interactions.

In my second study, my co-author Rahul Sami and I conducted human-subject laboratory experiments on variants of market scoring rule prediction markets, under different information distribution patterns, in order to evaluate the efficiency and speed of information aggregation, as well as test recent theoretical results on manipulative behavior by traders. We find that markets structured to have a fixed sequence of trades exhibit greater accuracy of information aggregation than the typical form that has unstructured trade. Prior theoretical predictions of differing strategic behavior under complementary information distributions and substitute information distributions are confirmed when the trading order is structured, but not in markets with an unstructured trading order. In the case of the markets with a structured order, we find that the information aggregation is consequently slower when information is complementary, as traders more frequently engage in bluffing and delaying strategies. In comparing two commonly used mechanisms, we find no significant difference between the performance of the direct probability-report form and the indirect security-trading form of the market scoring rule.

In my third study, I use game theory to analyze the performance of two mechanisms: the minimum threshold mechanism, under which one can only access the public goods if her contribution is higher than a pre-specified threshold, and the ratio mechanism, under which a user consumes at most an amount proportional to her own contribution level. I derive equilibrium predictions for these two mechanisms and analyze their performance in terms of social welfare. My results indicate some advantages of the minimum threshold mechanism over the ratio mechanism. There exist some conditions under which the minimum threshold mechanism can achieve the social optimum, but the ratio mechanism cannot. Furthermore, if the ratio mechanism implements a no-exclusion equilibrium, the same outcome can always be implemented by the minimum threshold mechanism.

These three studies are among the first steps towards systematically understanding strategic behavior of individuals in information aggregation systems. Much future work is needed in a number of directions to help better motivate high quality information provision to information aggregation systems. Here I highlight a few example research problems:

First, in market-based information systems, such as prediction markets, there is much to learn about users' potential manipulative behaviors. Theories in this area have been developed while making a few unrealistic assumptions, such as traders arriving at the market following a pre-specified order, traders all being risk neutral etc. Theories need to be developed to better predict users' manipulative behaviors while relaxing these assumptions, and study how these factors might affect the aggregated performance of prediction markets.

Second, to extend my work in chapter 4, I plan to design incentives for contributions that take into account contributor diversity. Information contribution is different from monetary donation. Money is substitutable — my contribution of $100 is just as good as your $100 — but information is often not. For some information systems, such as opinion forums, having diverse perspectives is just as important as having high participation.

Last, the growing practice of incorporating social networks in information systems presents new opportunities for motivating information contribution. It might bring changes to contributors' preferences — one might care differently about being a good friend than about being a good contributor. How might the presence of social networks in the information systems affect contribution?

Appendices

Appendix A

Results of robustness tests and simulations for Chapter 2

A.1 Robustness test of the timing distribution functional form assumptions

In Section 2.5 we reported our results estimated under the assumption that feedback timing follows a lognormal distribution. To evaluate the robustness our results to the assumed functional form of the distribution, we also estimated the model with Weibull and Gamma distributions and report our results in Table A.1.[1] For the lognormal function, we report the estimated coefficients of the independent variables on the four dependent variables, B_y, B_r, S_y, and S_r, and their associated p-values for two-sided tests that they are not different from zero. To compare the results estimated using Weibull and Gamma distributions to the lognormal distribution, we report the percentage change (under the column "% Δ") in each coefficient or parameter, that is, the percentage difference from the benchmark coefficient estimate under the lognormal assumption.

From Table A.1 we conclude that the specific functional form assumption does not affect our results materially. Most percentage changes in the coefficients are under 10%, except for three (highlighted in bold), all of which are below 25%.

[1] Based on our priors, we only considered distributions with non-negative support, and which can be asymmetric with a long tail.

		Lognormal		Gamma	Weibull
		Coef.	**P-Value**	**% Err.**	**% Err.**
B_y	**newbuyer**	-0.11	-6.04	-0.48	-0.70
	newseller	-0.28	-11.78	-0.49	-0.20
	newint	-0.04	-1.5	0.43	2.32
	lnfbbuyerr	0.19	37.54	-0.16	-0.02
	lnfbsellerr	-0.03	-8.38	0.44	0.50
	lnprice	0.04	13.48	-1.56	-0.64
	intercept	-0.48	-19.2	0.94	0.31
B_r	**newbuyer**	0.18	4.91	2.75	0.41
	newseller	-0.29	-6.16	1.00	-1.14
	newint	0.01	0.16	**-20.25**	**-15.91**
	lnfbbuyerr	0.10	9.67	-0.66	-2.51
	lnfbsellerr	-0.07	-11.64	-0.85	-1.22
	lnprice	-0.05	-7.29	0.39	-0.44
	intercept	-0.31	-6.48	-4.69	0.46
S_y	**newbuyer**	-0.28	-7.89	0.32	1.15
	newseller	-0.09	-2.54	0.43	0.51
	newint	-0.12	-2.95	0.07	-0.68
	lnfbbuyerr	0.08	10.45	0.47	1.40
	lnfbsellerr	0.20	27.61	-0.04	-0.56
	lnprice	-0.04	-8.02	-0.03	2.24
	intercept	-0.27	-6.19	-0.52	-2.60
S_r	**newbuyer**	-0.31	-4	2.66	3.21
	newseller	0.49	7.13	-0.48	3.48
	newint	0.09	1.03	5.70	8.42
	lnfbbuyerr	0.03	2.07	**10.62**	7.43
	lnfbsellerr	0.20	13.87	0.25	0.85
	lnprice	0.10	8.7	-1.58	3.46
	intercept	-1.60	-17.26	-0.88	-4.44

Table A.1: Coefficient robustness under different timing distribution assumptions. "% Δ" indicates the percentage difference from the benchmark coefficient estimate. Changes of more than 20% highlighted in bold.

A.2 Robustness tests for the definition of "new trader"

We defined a "new" trader to be one with fewer than 5 feedbacks. In Table A.2 we report the estimates obtained using different cutoffs, i.e., 3, 7, and 10 feedbacks. Most of the percentage errors are reasonably low, i.e., less than 50%. A few highlighted percentage errors in these are more than 50%. Some of the estimated coefficients that vary greatly across different cutoff values tend to have high p-values. For instance, the coefficient on *newint* for the dependent variable B_r has a percentage error of 1026.02% when estimated with Cutoff = 3. But the benchmark estimate of it with Cutoff = 5 has a p-value of 0.87. Thus this estimate was not statistically different from zero in the benchmark estimate. Examining other estimates with high percentage errors reveals that they do not alter our results qualitatively, though the size of the marginal effects might vary significantly depending on the cutoff values. This result reinforces our conclusion: "newness" matters.

A.3 Simulation results

To validate our maximum likelihood model, we conducted Monte Carlo simulations. We report the simulation parameters and results in Table A.3. Our simulated sample has 959,657 data points, the same sample size as our actual dataset.

First, we arbitrarily picked a set of "true" parameter values, as shown in the "True Val" column, as the targeted true values to estimate. Next, we randomly generated the following independent variables each using a uniform distribution on $[0, 1]$: *newbuyer*, *newseller*, *newint*, $lnfbbuyerr$, $lnfbsellerr$, and *lnprice*. Using the true coefficient values specified by us and the simulated independent variables, we generated the probabilities of strategy choices for each trader, i.e., either B_y and B_r or S_y and S_r depending on the role of the trader. We then use these probabilities to randomly select the "actual" strategy that the trader used. Last, using the selected strategies, and the parameters we specified for the timing functions, we generated a time stamp for each feedback given, if according to the trader's strategy she would give a feedback. Taking this simulated dataset, we then estimated our model with lognormal timing distributions, to obtain the coefficients shown in the "Coef" column. For each estimated coefficient, we also report its standard error (in the "SE" column), and the p-value of a two-sided test that the coefficient is not different from

		Cutoff 5		Cutoff 3	Cutoff 7	Cutoff 10
		Coef.	**P-Value**	**% Err.**	**% Err.**	**% Err.**
B_y	**newbuyer**	-0.11	0.00	**-110.86**	-19.35	-7.82
	newseller	-0.28	0.00	7.54	24.36	13.77
	newint	-0.04	0.13	-34.93	-32.40	23.22
	lnfbbuyerr	0.19	0.00	-9.55	-2.26	-13.82
	lnfbsellerr	-0.03	0.00	14.16	8.32	-18.15
	lnprice	0.04	0.00	3.71	3.23	-1.21
	intercept	-0.48	0.00	8.23	2.40	28.01
B_r	**newbuyer**	0.18	0.00	-14.80	**-77.24**	**52.70**
	newseller	-0.29	0.00	9.38	19.16	-11.88
	newint	0.01	0.87	**1026.02**	**-258.55**	**392.86**
	lnfbbuyerr	0.10	0.00	-9.89	-37.60	23.65
	lnfbsellerr	-0.07	0.00	18.75	-5.65	-14.23
	lnprice	-0.05	0.00	-0.74	-0.35	18.68
	intercept	-0.31	0.00	-12.81	42.07	-18.19
S_y	**newbuyer**	-0.28	0.00	-10.18	27.41	11.87
	newseller	-0.09	0.01	**-222.20**	**57.76**	**209.76**
	newint	-0.12	0.00	-24.01	20.49	32.42
	lnfbbuyerr	0.08	0.00	12.62	14.32	-21.85
	lnfbsellerr	0.20	0.00	-12.09	3.88	4.12
	lnprice	-0.04	0.00	5.93	-0.43	-15.40
	intercept	-0.27	0.00	19.03	-26.86	23.56
S_r	**newbuyer**	-0.31	0.00	13.44	33.41	24.05
	newseller	0.49	0.00	-31.34	-20.20	15.88
	newint	0.09	0.30	5.40	**-60.22**	**-79.11**
	lnfbbuyerr	0.03	0.04	**62.42**	47.21	-39.76
	lnfbsellerr	0.20	0.00	-22.18	-5.95	9.00
	lnprice	0.10	0.00	0.72	-0.19	-2.87
	intercept	-1.60	0.00	6.61	-0.51	-1.26

Table A.2: Robustness test result on the number of feedbacks that defines new trader. "% Δ" indicates the percentage difference from the benchmark coefficient estimate. Changes of more than 50% highlighted in bold.

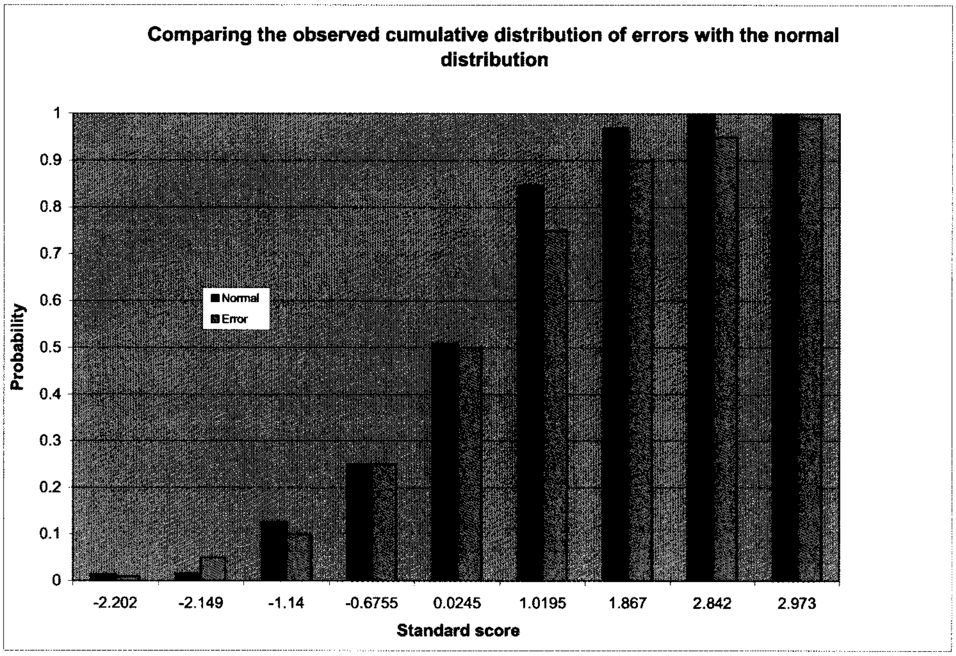

Figure A.1: Comparing the cumulative distribution of the simulation errors with the normal distribution.

zero (in the "P-val" column). To analyze the simulation error, we compute the deviation measured as the number of standard deviations, and report this in the final column.

Overall, our simulation results indicate that our model is valid. With approximately one million random draws, the distribution of simulation errors should converge close to a normal distribution. In Figure A.1 we plot the cumulative distribution of the normalized simulation errors against the predicted normal error distribution. The tails are a bit heavier, but overall the fit is quite good and there are no outliers.

Depd Var	Indp Var	True Val	Coef	SE	P-val	Err in Std
B_y	newbuyer	0.000	0.042	0.015	0.004	2.842
	newseller	-0.250	-0.256	0.015	0.000	-0.422
	newint	0.150	0.133	0.015	0.000	-1.140
	lnfbbuyerr	-0.300	-0.309	0.015	0.000	-0.589
	lnfbsellerr	0.510	0.525	0.015	0.000	1.023
	lnprice	-0.100	-0.106	0.015	0.000	-0.400
	intercept	0.300	0.288	0.018	0.000	-0.669
B_r	newbuyer	0.000	0.067	0.031	0.029	2.177
	newseller	-0.550	-0.548	0.031	0.000	0.057
	newint	0.520	0.501	0.031	0.000	-0.608
	lnfbbuyerr	-0.200	-0.226	0.031	0.000	-0.842
	lnfbsellerr	0.150	0.182	0.031	0.000	1.029
	lnprice	0.300	0.279	0.031	0.000	-0.682
	intercept	-0.300	-0.323	0.038	0.000	-0.621
S_y	newbuyer	0.000	0.024	0.014	0.079	1.756
	newseller	-0.300	-0.311	0.014	0.000	-0.822
	newint	-0.350	-0.345	0.014	0.000	0.330
	lnfbbuyerr	-0.500	-0.486	0.014	0.000	1.016
	lnfbsellerr	0.350	0.353	0.014	0.000	0.232
	lnprice	-0.350	-0.324	0.014	0.000	1.867
	intercept	-0.400	-0.437	0.017	0.000	-2.202
S_r	newbuyer	0.000	0.010	0.012	0.418	0.811
	newseller	-0.200	-0.198	0.012	0.000	0.149
	newint	0.800	0.800	0.013	0.000	0.016
	lnfbbuyerr	-0.800	-0.763	0.012	0.000	2.973
	lnfbsellerr	0.450	0.423	0.012	0.000	-2.149
	lnprice	0.100	0.119	0.012	0.000	1.515
	intercept	0.400	0.372	0.016	0.000	-1.797
Timing Parameters	μ_{sy}	3.500	3.496	0.007	0.000	-0.542
	σ_{sy}	2.000	1.997	0.005	0.000	-0.698
	μ_{by}	4.000	4.000	0.003	0.000	0.033
	σ_{by}	2.000	2.001	0.002	0.000	0.487

Table A.3: Simulation Results with a sample of 959,657 data points.

Appendix B

The computer interfaces of the experimental prediction markets in Chapter 3

B.1 Computer Interfaces

Figures B.1 and B.2 show the trading interfaces used in our direct and indirect MSR markets respectively. They each contain two boxes: a probability or price update graph that displays the market activities in real time and a prediction or transaction submission box in which they enter their trades. With the direct MSR as shown in Figure B.1, all a trader has to do is to report her prediction in the box next to "My next prediction". In indirect MSR markets as shown in Figure B.2, traders look at the current prices for the black and white securities and make their trading decisions by clicking the trading buttons provided to them.

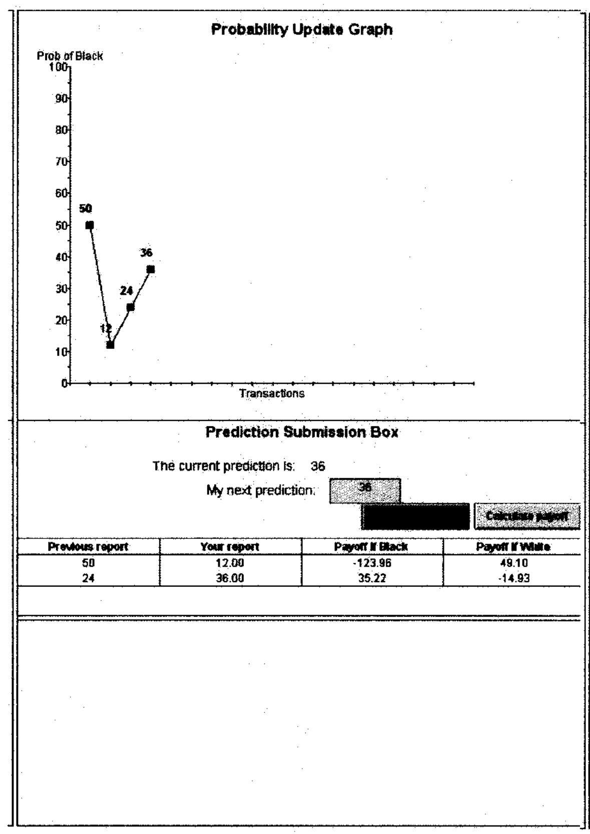

Figure B.1: The trading interface of the direct MSR markets

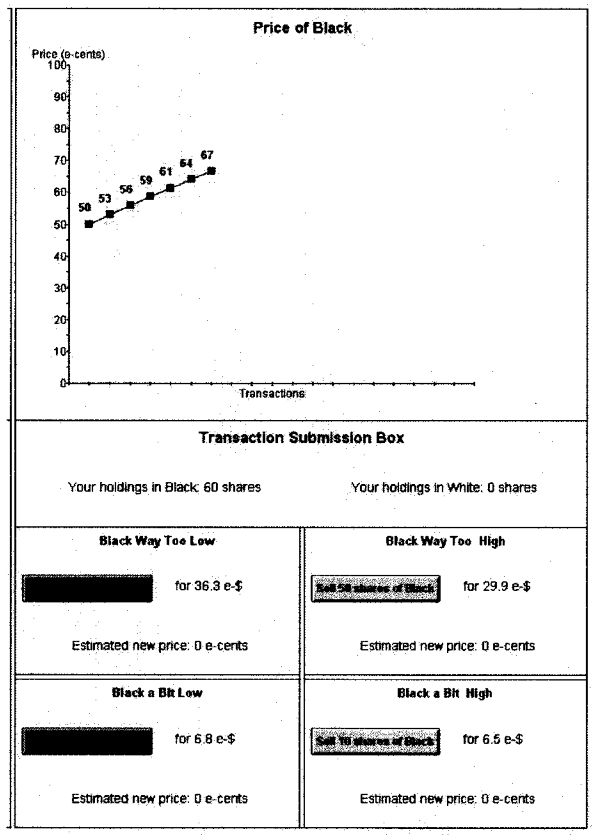

Figure B.2: The trading interface of the indirect MSR markets

Appendix C

Proofs of the facts, propositions, and lemmas in Chapter 4

C.1 Proof of Fact 1

Proof. In a Nash equilibrium, suppose the equilibrium level of public good production is y^{VCM}. Each individual is either a free-rider ($x_i = 0$) or a contributor who contributes an amount of x_i^* that maximizes her payoff. By setting the first-order condition to zero, we obtain the following condition that must hold for an arbitrary contributor i: $\alpha_i v'(y^{VCM}) = c'(x_i^*)$. Given that c' is an increasing function ($c'' > 0$), we know for all contributors, x^* increases in α. For free-riders, let i be a free-rider that contributes zero and we know $\alpha_i v'(y^{VCM}) \leq c'(0)$. It is obvious that any individual j with $\alpha_j > \alpha_i$ will be contributing either zero or a positive amount. □

C.2 Proof of Fact 2

Proof. Let $X = \{x \in \mathbb{R}^N : 0 \leq x_i \leq \omega \text{ for } i = 1, \cdots, N\}$ be the strategy space of the N individuals, which is nonempty, convex, and a compact subset of some Euclidean space $\mathbb{R}^N$. And we know that any individual, i, has a continuous and quasi-concave utility function: $u_i(x_i) = \alpha_i v(x_{-i} + x_i) + \omega - c(x_i)$. According to Proposition 8.D.3 in Mas-Colell et al. (1995), there exists a pure strategy Nash equilibrium.

I next establish the uniqueness of the pure strategy Nash equilibrium in two steps. First, I establish that given an equilibrium level of production, denoted as Y, there is a unique vector of contributions, x, with individual i contributing x_i. Next I show that for any profile of α ($\alpha = \{\alpha_1, \cdots, \alpha_N\}$), there exists a unique equilibrium level of production.

First, given Y, in equilibrium each individual i contributes x_i such that it maximizes her payoff. She sets $\alpha_i v'(Y) - c'(x_i) = 0$ (One can easily verify that the second order derivative is less than zero.). This first order condition determines a unique x_i for individual i.

Next, I establish that there cannot be more than one equilibrium levels of production, given a particular profile of α. Suppose there are two different equilibrium levels of production, Y and Y'. Without loss of generality, let $Y' > Y$. Let individual i be an individual who contributes a non-zero amount in both equilibria. Such an individual must exist due to the monotonicity in participants' contribution as shown in Fact 1. In the Y equilibrium individual i either contributes x_i^* where x_i^* satisfies the following equation:

$$\alpha_i v'(Y) - c'(x_i^*) = 0 \tag{C.1}$$

Similarly, in the Y' equilibrium, the same individual i contributes $x_i'^*$:

$$\alpha_i v'(Y') - c'(x_i'^*) = 0 \tag{C.2}$$

$Y' > Y$ implies $v'(Y') < v'(Y)$ (due to $v'' < 0$). Comparing equation (C.1) and (C.2), and by the assumption that $c'' > 0$, we know $x_i^* > x_i'^*$. Thus if an individual is a contributor under both equilibria, she must be contributing more under the Y equilibrium than under the Y' equilibrium.

Now, if we can show that there are also more contributors under the Y equilibrium than under the Y' equilibrium, we would have shown that it is impossible for Y' to be greater than Y. Let α_0 and α_0' denote the contributors with the lowest non-zero contribution in each equilibrium. They both contribute one unit of public goods under their respective equilibria. Applying the first order conditions as exemplified in equation (C.1) and (C.2) to these two contributors, we find that $\alpha_0 < \alpha_0'$. Given that x_i increases in α_i (due to Fact 1), the fact that $\alpha_0 < \alpha_0'$ means there are more contributors in the Y equilibrium than in the Y' equilibrium.

□

C.3 Proof of Fact 3

Proof. In a NE, individual i best responses to the equilibrium level of public good production, y^{MTM}. First define two terms, α_t and α_m, which satisfy the following two conditions respectively:

$$\alpha_t v(y^{MTM}) - c(t) = 0 \tag{C.3}$$

$$\alpha_m v'(y^{MTM}) - c'(t) = 0 \tag{C.4}$$

If $\alpha_i \geq \alpha_t$ and $\alpha_i \geq \alpha_m$, individual i chooses $x_i \geq t$ to best response to x_{-i}. In particular, $x_i = x_i^* = \arg\max_x \alpha_i v(x + x_{-i}) - c(x)$. If $\alpha_i \leq \alpha_t$ and $\alpha_i \leq \alpha_m$, $x_i = 0$. If $\alpha_i \geq \alpha_t$ and $\alpha_i \leq \alpha_m$, $x_i = t$. The last combination — $\alpha_i \leq \alpha_t$ and $\alpha_i \geq \alpha_m$ — is impossible, because $\alpha_t \leq \alpha_m$, as shown below.

First of all, for any arbitrary individual i producing the public goods standalone, we define two points: $\tilde{x}_i$ as individual i's standalone break-even point, and x_i^0 as individual i's standalone profit maximization point. By definition, $\tilde{x}_i$ satisfies the condition $u_i(\tilde{x}_i) = \alpha_i v(\tilde{x}_i) - c(\tilde{x}_i) = 0$, and x_i^0 satisfies $u_i'(x_i^0) = \alpha_i v'(x_i^0) - c'(x_i^0) = 0$ (due to the concavity of v and the convexity of c).

With these two points defined, the outline of the rest of the proof is as follows: I first show that x_i^0 increases in α_i; Then I show that $x_t^0 < x_m^0$, where x_t^0 and x_m^0 are the standalone profit maximization points of the individuals with α_t and α_m respectively. These two steps combined would immediately imply that $\alpha_t < \alpha_m$.

I first show that x_i^0 weakly increases in α_i. If $x_i^0 > 0$, we know $\alpha_i = \frac{c'(x_i^0)}{v'(x_i^0)}$. Because c' is an increasing function and v' is a decreasing function, we know x_i^0 increases in α_i. If $x_i^0 = 0$, then for any individual j with $\alpha_j > \alpha_i$, it is straightforward to show that $x_j^0 \geq x_i^0$.

My last step is to show $x_t^0 < x_m^0$. Given that both v and c are continuous functions, with $v(0) = c(0) = 0$, $v'(0) > c'(0) > 0$, $v'' < 0$ and $c'' > 0$, for each individual i, the functions $\alpha_i v(x)$ and $c(x)$ must intersect at a unique point. This point is $\tilde{x}_i$, by definition. If individual i produces any amount $x < \tilde{x}_i$, then we know $u_i(x) > 0$ and vice versa; and if she produces any amount $x > \tilde{x}_i$, then we know $u_i(x) < 0$ and vice versa. Since $u_i(x)$ is a continuous function, has derivative at each point of the open interval $(0, \tilde{x}_i)$, and $u_i(0) = u_i(\tilde{x}_i)$, by Rolle's theorem, there must be at least one point in the interval $(0, \tilde{x}_i)$ such that $u_i'(x) = 0$. This point is unique (due to the concavity of $u_i(x)$) and thus it has to be our x_i^0. Since this point is in the interval $(0, \tilde{x}_i)$, we know $x_i^0 < \tilde{x}_i$. In the interval $[0, x_i^0)$, $u_i' > 0$ and in the interval $(x_i^0, +\infty)$, $u_i' < 0$.

Equation (C.3) implies $\alpha_t v(t) - c(t) < 0$ (due to $t < y^{MTM}$), which implies $t > \tilde{x}_i$ due to the characterization of $\tilde{x}_i$ in the previous paragraph. Since I have shown above that $x_i^0 < \tilde{x}_i$ for any individual i, we know $t > x_t^0$. Equation (C.4) implies $\alpha_m v'(t) - c'(t) > 0$ (due to $t < y^{MTM}$ and $v'' < 0$), which implies $t < x_m^0$ due to the characterization of x_i^0 in the previous paragraph. Combining $t > x_t^0$ and $t < x_m^0$, we obtain $x_t^0 < x_m^0$.

□

C.4 Proof of Fact 4

Proof. I complete the proof of Fact 4 in two steps. In step 1 I show that there exists an MTM that strictly increases the total amount of public good production compared to VCM. In Step 2, I show that if the total amount of production under this MTM equilibrium is not higher than the socially optimal amount, this MTM equilibrium must be welfare improving compared to VCM.

Step 1. Suppose in the equilibrium under VCM, the total amount of public goods produced is y^{VCM}. Let us pick a t as the threshold for the MTM, such that $\underline{\alpha} v(y^{VCM}) - c(t) = 0$. I first show that such a t always exists. Next, I establish that under such an MTM, there exists a NE such that no one is excluded (everyone contributes at least t), and the total amount of public goods produced, y^{MTM}, is greater than y^{VCM}.

First, I show that we can always find a t such that $\underline{\alpha} v(y^{VCM}) - c(t) = 0$. Suppose under VCM, the individual with $\underline{\alpha}$ contributes zero. Since $v(0) = 0$ and $v' > 0$, we know that $\underline{\alpha} v(y^{VCM}) > 0$ ($y^{VCM} > 0$ by assumption). Given that $c(0) = 0$ and $c' > 0$, we can always find a $t > 0$ such that $\underline{\alpha} v(y^{VCM}) - c(t) = 0$. Suppose under VCM, the individual with $\underline{\alpha}$ contributes $x_{\underline{\alpha}}^{VCM}$ ($x_{\underline{\alpha}}^{VCM} > 0$). In equilibrium, $\underline{\alpha} v'(y^{VCM}) - c'(x_{\underline{\alpha}}^{VCM}) = 0$, which implies $\underline{\alpha} v'(x_{\underline{\alpha}}^{VCM}) - c'(x_{\underline{\alpha}}^{VCM}) > 0$ (due to $x_{\underline{\alpha}}^{VCM} < y^{VCM}$ and $v'' < 0$). Let $p(x) = \underline{\alpha} v(x) - c(x)$. By our assumptions of the v and c functions, we know $p(0) = 0$, $p'(0) > 0$, and p is a concave function. Thus $p'(x_{\underline{\alpha}}^{VCM}) > 0$ and $x_{\underline{\alpha}}^{VCM} > 0$ imply $p(x_{\underline{\alpha}}^{VCM}) = \underline{\alpha} v(x_{\underline{\alpha}}^{VCM}) - c(x_{\underline{\alpha}}^{VCM}) > 0$, which further implies $\underline{\alpha} v(y^{VCM}) - c(x_{\underline{\alpha}}^{VCM}) > 0$ (due to $y^{VCM} > x_{\underline{\alpha}}^{VCM}$ and $v' > 0$). We then must be able to find a $t > x_{\underline{\alpha}}^{VCM}$ such that $\underline{\alpha} v(y^{VCM}) - c(t) = 0$.

Given such a t, and suppose the level of public good production, y^{MTM}, is greater than y^{VCM}, I next show that no one will deviate from the NE equilibrium. That is, everyone will contribute at least t. Since $\underline{\alpha} v(y^{VCM}) - c(t) = 0$, we know the individuals with $\underline{\alpha}$ will contribute exactly t. Everyone else with a value coefficient greater than $\underline{\alpha}$ will have a positive utility if they contribute t. Thus no one will deviate from contributing at least t.

To prove the existence of a NE with $y^{MTM} > y^{VCM}$, we need to show that in an equilibrium with everyone contributing at least t, the total amount of production, y^{MTM}, is indeed greater than y^{VCM}. Now, suppose $y^{MTM} < y^{VCM}$. Let us further suppose that under VCM, for any $\alpha_i \in [\alpha_k, \bar{\alpha}]$, $x_i^{VCM} > t$, and for any $\alpha_j \in [\underline{\alpha}, \alpha_k)$, $x_j^{VCM} \leq t$. For individual i, we know $\alpha_i v'(y^{VCM}) - c'(x_i^{VCM}) = 0$ under VCM, and $\alpha_i v'(y^{MTM}) - c'(x_i^{MTM}) = 0$ under MTM, where x_i^{MTM} is individual i's contribution level in the NE under MTM. Since $y^{MTM} < y^{VCM}$, we know $x_i^{MTM} > x_i^{VCM}$. For individual j, we know $x_j^{MTM} \geq t \geq x_j^{VCM}$. Further, from the above analysis we know that the threshold t was set such that the individual with $\underline{\alpha}$ contributes strictly less than t under the VCM equilibrium. Since in the MTM equilibrium, everyone contributes weakly more than in the VCM equilibrium, and at least some individuals, with $\underline{\alpha}$, contribute strictly more in the MTM equilibrium, it contradicts our initial assumption that $y^{MTM} < y^{VCM}$.

Step 2. I show that if y^{MTM} is not higher than the socially optimal amount, this equilibrium must be creating higher welfare than the VCM equilibrium. I have shown above that we can always find a t such that $y^{MTM} > y^{VCM}$. Let us restrict the t such that the resulting y^{MTM} is not higher than the socially optimal amount. Again, let α_k divide the participants who contribute less than t and more than t under the VCM equilibrium. Compared to VCM, under MTM, individuals with $\alpha \in [\underline{\alpha}, \alpha_k]$ (weakly) increase their contribution to t, whereas individuals with $\alpha \in (\alpha_k, \bar{\alpha}]$ (weakly) decrease their contribution (but still contribute at least t). We can see the latter by focusing on an arbitrary individual j with $\alpha_j > \alpha_k$. We know $\alpha_j v'(y^{VCM}) - c'(x_j^{VCM}) = 0$ and $\alpha_j v'(y^{MTM}) - c'(x_j^{MTM}) = 0$ under the VCM and MTM equilibria respectively. Since $y^{MTM} > y^{VCM}$, it must be that $x_j^{MTM} < x_j^{VCM}$ (due to $v'' < 0$).

Compared to VCM, let the total amount of increased production under MTM by the individuals with $\alpha \in [\underline{\alpha}, \alpha_k]$ be Γ:

$$\Gamma = \int_{\underline{\alpha}}^{\alpha_k} (x_\alpha^{MTM} - x_\alpha^{VCM}) f(\alpha) \mathrm{d}\alpha, \tag{C.5}$$

Compared to VCM, let the total amount of decreased production under MTM by the individuals with $\alpha \in (\alpha_k, \bar{\alpha}]$ be Λ:

$$\Lambda = \int_{\alpha_k}^{\bar{\alpha}} (x_\alpha^{VCM} - x_\alpha^{MTM}) f(\alpha) \mathrm{d}\alpha. \tag{C.6}$$

We know $\Gamma > 0$ and $\Lambda > 0$. Let Δ be the difference, thus $\Delta = \Gamma - \Lambda$. We know from the above analysis that $\Delta > 0$. Based on these definitions, we can divide Γ into two portions, one with the amount of Λ and the other with the amount of Δ. Under VCM, this Λ amount is produced by the participants with $\alpha > \alpha_k$. Under MTM, a same amount is produced by participants with $\alpha \leq \alpha_k$. As the cost function $c(x)$ is convex, the total cost of producing Λ must be lower under MTM than under VCM, because the contributors under VCM face higher marginal costs than the contributors under MTM. As a result, producing this Λ amount of public goods under MTM provides higher social welfare than under VCM.

Now we only need to show that the Δ portion creates more utility for all the participants than it costs. We know that y^{MTM} is not higher than the socially optimal amount, y^{FB}, so y^{MTM} must satisfy the following condition:

$$\left(N \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha) \mathrm{d}\alpha\right) v'(y^{MTM}) > c'(\frac{y^{FB}}{N}) \tag{C.7}$$

The Δ amount is produced by the participants with $\alpha < \alpha_k$, who each contributes t in the equilibrium under MTM. Given that in an equilibrium with a total amount of production y^{MTM}, all the participants contribute at least t, we know $t \leq \frac{y^{MTM}}{N}$, hence $t \leq \frac{y^{FB}}{N}$. Since $c'' > 0$, we know $c'(t) < c'(\frac{y^{FB}}{N})$. By equation (C.7), we know that $\left(N \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha) \mathrm{d}\alpha\right) v'(y^{MTM}) > c'(t)$. That is, producing the extra Δ amount of public goods under MTM leads to a welfare increase. $\square$

C.5 Proof of Proposition 1

Proof. We know that in a socially optimal outcome, everybody contributes y^{FB}/N, and y^{FB} satisfies $N \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha) \mathrm{d}\alpha v'(y^{FB}) = c'(\frac{y^{FB}}{N})$. Let us define a p function by $p(y) = \frac{c'(\frac{y}{N})}{v'(y)}$. Since c' is a monotonically increasing function and v' is a monotonically decreasing function, we know p is a monotonically increasing function. Thus the inverse function of p, p^{-1}, exists and it monotonically increases. We can then write $y^{FB} = p^{-1}(N \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha) \mathrm{d}\alpha)$.

A necessary condition for MTM to achieve the socially optimal outcome is that the individual with $\underline{\alpha}$ would not deviate from this socially optimal outcome, implying that equation (4.3) should hold. It is also a sufficient condition because if the individual with $\underline{\alpha}$ would be willing to contribute at least t, everybody else would also be willing to contribute at least t. In addition, no one would profit from contributing more than t, which can be demonstrated by individual $\bar{\alpha}$'s profit maximization problem. The first order derivative of individual $\bar{\alpha}$'s payoff function is less than zero: $\bar{\alpha}v'(y^{FB}) - c'(y^{FB}/N) < 0$ (due to $\bar{\alpha} \leq N\int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha$). That is, if equation (4.3) holds, everyone would contribute exactly y^{FB}/N.

□

C.6 Proof of Proposition 2

Proof. Next, I show that if $v'(y) + yv''(y) < 0$ for all $y > 0$, the right hand side of equation (4.3) decreases in N. Given the assumption that $F(\alpha)$ is stationary as N increases, let $\Omega = \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha f(\alpha)\mathrm{d}\alpha$ and take the first order derivative of the right hand side of equation (4.3) with respect to N. We obtain the following:

$$\frac{\partial \frac{c(\frac{1}{N}p^{-1}(N\Omega))}{v(p^{-1}(N\Omega))}}{\partial N} = \frac{c'(p^{-1}(N\Omega)/N)\frac{\partial(p^{-1}(N\Omega)/N)}{\partial N}}{v(p^{-1}(N\Omega))} - \frac{c(p^{-1}(N\Omega)/N)v'(p^{-1}(N\Omega))}{v^2(p^{-1}(N\Omega))}\frac{\partial p^{-1}(N\Omega)}{\partial N}$$

The sign of the above partial derivative hinges on the sign of the term $\frac{\partial(p^{-1}(N\Omega)/N)}{\partial N}$, because $v > 0$, $c > 0$, $v' > 0$, $c' > 0$, and it is easy to verify that $\frac{\partial p^{-1}(N\Omega)}{\partial N} > 0$. As $p^{-1}(N\Omega)/N$ is the level of individual contribution in the socially optimal outcome, $\frac{\partial(p^{-1}(N\Omega)/N)}{\partial N} = \frac{\partial x^{FB}}{\partial N}$, where x^{FB} is the level of individual contribution in the socially optimal outcome. Taking the first order derivative with respect to N on both sides of the equation, $N\Omega v'(Nx^{FB}) - c'(x^{FB}) = 0$, we arrive at

$$\frac{\partial x^{FB}}{\partial N} = \Omega\frac{v'(Nx^{FB}) + Nx^{FB}v''(Nx^{FB})}{c''(x^{FB}) - N^2\Omega v''(Nx^{FB})} \tag{C.8}$$

Obviously, the condition $v'(Nx^{FB}) + Nx^{FB}v''(Nx^{FB}) < 0$ would imply $\frac{\partial x^{FB}}{\partial N} < 0$.

□

C.7 Proof of Proposition 3

Proof. The existence of a pure strategy Nash equilibrium follows from Proposition 8.D.3 in Mas-Colell et al. (1995). In a Nash equilibrium, individual i chooses a contribution level, x_i, to best response to x_{-i}, the total amount of public goods produced by all other participants. The ratio mechanism requires that an individual, i, contributing x_i, has access to an amount of public goods as determined by $min\{rx_i, x_{-i} + x_i\}$.

Let us use x_i^r to denote individual i's r-optimal amount of contribution. Individual i's best response function is as below:

$$x_i = \begin{cases} \arg\max_x \alpha_i v(x_{-i}+x) - c(x) & \text{if } x_{-i} < x^*_{-i} \\ x_{-i}/(r-1) & \text{if } x^*_{-i} \leq x_{-i} < (r-1)x_i^r \\ x_i^r & \text{if } x_{-i} \geq (r-1)x_i^r \end{cases}$$

where x^*_{-i} is a parameter (independent of others' contribution levels) for individual i that satisfies the following condition: $\alpha_i v'(\frac{rx^*_{-i}}{r-1}) - c'(\frac{x^*_{-i}}{r-1}) = 0$. In a Nash equilibrium, suppose y^* is the total amount of public good production. Let individual k with α_k be the participant with x_k^r such that $rx_k^r = x^*_{-k} + x_k^r$, which is equivalent to $x_k^r = x^*_{-k}/(r-1)$. According to the best response function we derived, in equilibrium, an individual i with an $\alpha_i \leq \alpha_k$ will contribute x_i^r.

Any individual j with an $\alpha_j \in (\alpha_k, \bar{\alpha}]$ maximizes her utility by contributing $\arg\max_x \alpha_j v(x_{-j}+x) - c(x)$ where $x \geq x_j^r$. In equilibrium, if $\alpha_j v'(y^*) - c'(y^*/r) < 0$, individual j maximizes her utility by contributing y^*/r. If $\alpha_j v'(y^*) - c'(y^*/r) > 0$, she would then contribute an amount greater than y^*/r. We find for participant k, $\alpha_k r v'(rx_k^r) - c'(x_k^r) = 0$, implying $\alpha_k v'(y^*) - c'(x_k^r) < 0$. Thus participant k will maximize her utility by contributing $x_k^r = y^*/r$.

Let us then focus on the individuals with their value coefficients greater than α_k. Let m be the marginal individual who is indifferent between contributing y^*/r and $y^*/r + \epsilon$, where ϵ is an infinitesimally small amount. Thus α_m satisfies the following condition: $\alpha_m v'(y^*) - c'(y^*/r) = 0$. Since in equilibrium x_i^* weakly increases in α_i, all the individuals with an α greater than α_m will contribute more than y^*/r. □

C.8 Proof of Lemma 1

Proof. $x_i^r = \arg\max_x \alpha_i v(rx) - c(x)$. Since $v'' < 0$ and $c'' > 0$, there is a global maximizer x_i^r which satisfies the following first order condition:

$$\alpha_i r v'(rx) - c'(x) = 0 \tag{C.9}$$

Taking partial derivatives of both sides of equation (C.9) we obtain

$$\frac{\partial x_i^r}{\partial r} = \alpha_i \frac{v'(rx_i^r) + rx_i^r v''(rx_i^r)}{c''(x_i^r) - \alpha_i r^2 v''(rx_i^r)}$$

Apparently, $v'(rx_i^r) + rx_i^r v''(rx_i^r)$ and $\frac{\partial x_i^r}{\partial r}$ have the same sign.

□

C.9 Proof of Proposition 5

Proof. Suppose under an RM Nash equilibrium everyone has full access to the public goods. Let y^* be the equilibrium amount of public good production. According to Proposition 3, there must be at most two types of participants: those who contribute exactly y^*/r and those who contribute more than y^*/r. For any individual i who contributes exactly y^*/r, the following condition must hold: $\alpha_i rv'(y^*) - c'(y^*/r) \geq 0$. This is the result of individual i's utility maximization, while contributing zero and receiving zero payoff is in her choice set. Given that individual i chooses to contribute more than zero amount ($y^* > 0$), it must be that $\alpha_i v(y^*) - c(y^*/r) \geq 0$, which implies if it is under a MTM with $t = y^*/r$, individual i must also be willing to contribute t in order to gain full access to the y^* amount of public goods. I next show that any individual j who contributes more than y^*/r amount under the RM equilibrium must find contributing the same amount maximizes her utility under the corresponding MTM equilibrium as well. Suppose individual j contributes x_j amount under the RM equilibrium, it must be that $\alpha_j v'(y^*) - c'(x_j) = 0$. The same condition would also hold under the MTM equilibrium with the same y^* level of total production. In summary, the same outcome in this RM non-exclusion equilibrium constitutes a Nash equilibrium as well under the MTM with $t = y^*/r$. □

Bibliography

B Thomas Adler and Luca de Alfaro. A content-driven reputation system for the wikipedia. In *Proceedings of the 16th international conference on World Wide Web*, pages 261–270, 2007.

Anat R. Admati and Motty Perry. Joint projects without commitment. *Review of Economic Studies*, Jan 1991.

George A. Akerlof. A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, 94(4):749–775, June 1980.

James Andreoni and John H. Miller. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal*, 103(418):570–585, May 1993.

Solomon E. Asch. Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70(416), 1956.

Sulin Ba and Paul A. Pavlou. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly*, 26(3): 243–268, September 2002.

Parimal Kanti Bag and Eyal Winter. Simple subscription mechanisms for excludable public goods. *Journal of Economic Theory*, 87:72–94, 1999. URL `http://ratio.huji.ac.il/dp/dp199.pdf`.

Mark Bagnoli and Barton L. Lipman. Provision of public goods: Fully implementing the core through private contributions. *Review of Economic Studies*, Jan 1989.

Henry Berg and Todd A Proebsting. Hanson's automated market maker. *The Journal of Prediction Markets*, (1):45–49, April 2009. URL `http://www.ingentaconnect.com/content/ubpl/jpm/2009/00000003/00000001/art00006`.

Joyce E. Berg, Forrest D. Nelson, and Thomas A. Rietz. Prediction market accuracy in the long run. *International Journal of Forecasting*, 24(2):285 – 300, 2008. URL `http://www.sciencedirect.com/science/article/B6V92-4SCTMTB-1/2/6360a13dc244ab81add6760d3e84584e`.

B. Douglas Bernheim. A theory of conformity. *The Journal of Political Economy*, 102(5):841–877, October 1994.

Gary Bolton, Ben Greiner, and Axel Ockenfels. Engineering trust - reciprocity in the production of reputation information. Working paper, 2009.

Luís Cabral and Ali Hortaçsu. The dynamics of seller reputation: Theory and evidence from eBay. *Journal of Industrial Economics*, 58:54–78, 2010.

Yan Chen. An experimental study of serial and average cost pricing mechanisms. *Journal of Public Economics*, 87:235–255, 2003.

Yiling Chen and David M. Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence (UAI 2007)*, pages 49–56, Vancouver, BC Canada, 2007.

Yiling Chen, Daniel M. Reeves, David M. Pennock, Robin D. Hanson, Lance Fortnow, and Rica Gonen. Bluffing and strategic reticence in prediction markets. In *Lecture Notes in Computer Science, The Third Workshop on Internet and Network Economics (WINE 2007)*, pages 70–81, 2007.

Yiling Chen, Stanko Dimitrov, Rahul Sami, Daniel Reeves, David Pennock, Robin Hanson, Lance Fortnow, and Rica Gonen. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, 2009.

Bram Cohen. Incentives build robustness in bittorrent. In *Workshop on Economics of Peer-to-Peer systems*, CA, June 2003. Berkeley. URL `http://www.ittc.ku.edu/~niehaus/classes/750-s05/documents/BT-description.pdf`.

Roger M. Cooke. *Experts in Uncertainty: Opinion and Subjective Probability in Science.* Oxford University Press, New York, 1991.

Bo Cowgill, Justin Wolfers, and Eric Zitzewitz. Using prediction markets to track information flow: Evidence from google, 2008. URL `http://www.bocowgill.com/GooglePredictionMarketPaper.pdf`.

Rajat Deb and Laura Razzolini. Auction-like mechanisms for pricing excludable public goods. *Journal of Economic Theory*, Jan 1999. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022053199926030`.

Chrysanthos Dellarocas. The digitization of word-of-mouth: promise and challenges of online reputation mechanisms. *Management Science*, 49(10):1407–1424, October 2003.

Chrysanthos Dellarocas and Charles A. Wood. The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science*, 54(3), March 2008.

Stanko Dimitrov and Rahul Sami. Non-myopic strategies in prediction markets. In *Proceedings of the Nineth ACM Conference on Electronic Commerce (EC'08)*, Chicago, IL, July 2008. ACM.

Stanko Dimitrov and Rahul Sami. Composition of markets with conflicting incentives. In *The Proceedings of the 2010 ACM Conference on Electronic Commerce (EC'10)*, Boston, MA, June 2010. ACM.

E. F. Fama. Efficient capital market: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

Ernst Fehr and Simon Gächter. Altruistic punishment in humans. *Nature*, 415(10), January 2002.

Michal Feldman, Christos Papadimitriou, John Chuang, and Ion Stoica. Free-riding and whitewashing in peer-to-peer systems. In *PINS '04: Proceedings of the ACM SIGCOMM workshop on Practice and theory of incentives in networked systems*, pages 228–236, New York, NY, USA, 2004. ACM Press.

Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.

Robert Forsythe, Forrest Nelson, George R. Neumann, and Jack Wright. Anatomy of an experimental political stock market. *American Economic Review*, 82(5), 1992.

Simon Gächter and Armin Falk. Reputation and reciprocity: Consequences for the labour relation. *Scandinavian Journal of Economics*, 104(1):1–26, 2002.

Sean Gailmard and Thomas R Palfrey. An experimental comparison of collective choice procedures for excludable public goods. *Journal of Public Economics*, 89: 1361–1398, Aug 2005.

Robert Gazzale. *Giving Gossips Their Due: Information Provision in Games with Private Monitoring*. PhD thesis, University of Michigan, Ann Arbor, MI, 2004.

Andreas Graefe and J. Scott Armstrong. Comparing face-to-face meetings, nominal groups, delphi and prediction markets on an estimation task, 2008. Working paper.

William Greene. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal*, 7: 98–119, 2004.

Theodore Groves and John Ledyard. Optimal allocation of public goods: A solution to the "free rider" problem. *Econometrica*, 45(4):783–809, May 1977. URL `http://links.jstor.org/sici?sici=0012-9682%28197705%2945%3A4%3C783%3AOAOPGA%3E2.0.CO%3B2-8`.

Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.

Robin Hanson, Ryan Oprea, and David Porter. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior and Organization*, Jan 2006. URL `http://linkinghub.elsevier.com/retrieve/pii/S0167268105001575`.

Friedrich A. Hayek. The use of knowledge in society. *The American Economic Review*, 35:519 – 530, 1945.

Paul J Healy, John O Ledyard, Sera Linardi, and J Richard Lowery. Information aggregation mechanisms and environmental complexity. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.130.7642&rep=rep1&type=pdf`. Working paper, 2009.

Lian Jian and Jeffrey K. MacKie-Mason. Why share in peer-to-peer networks? In *ICEC '08: Proceedings of the 10th international conference on Electronic commerce*, pages 1–8, New York, NY, USA, 2008. ACM.

Tapan Khopkar, Xin Li, and Paul Resnick. Self-selection, slipping, salvaging, slacking, and stoning: the impacts of negative feedback at eBay. In *EC '05: Proceedings of the 6th ACM conference on Electronic commerce*, pages 223–231, New York, NY, USA, 2005. ACM.

Cliff Lampe and Paul Resnick. Slash(dot) and burn: distributed moderation in a large online conversation space. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 543–550, 2004. URL `http://doi.acm.org/10.1145/985692.985761`.

John Ledyard, Robin Hanson, and Takashi Ishikida. An experimental test of combinatorial information markets. *Journal of Economic Behavior and Organization*, 69 (2):182–189, 2009.

Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

Herve Moulin. Serial cost-sharing of excludable public goods. *The Review of Economic Studies*, 61(2):305–325, apr 1994.

John F. Muth. Rational expectations and the theory of price movements. *Econometrica*, 29:315–335, 1961.

Ryan Oprea, David Porter, Chris Hibbert, Robin Hanson, and Dorina Tila. Can manipulators mislead market observers? *Working paper*, 2007.

David Pennock. Implementing Hanson's Market Maker, 2006. Blog post, available at `http://blog.oddhead.com/2006/10/30/implementing-hansons-market-maker/`.

Charles R. Plott and Shyam Sunder. Efficiency of experimental security markets with insider information: an application of rational-expectations models. *The Journal of Political Economy*, 90:663–698, 1982.

Charles R. Plott and Shyam Sunder. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometric*, 56(5):1085–1118, 1988.

Roy Radner. Rational expectations equilibrium: Generic existence and the information revealed in prices. *Econometrica*, 47(3):655–678, 1979.

Laura Razzolini, Michael Reksulak, and Robert Dorsey. An experimental evaluation of the serial cost sharing rule. *Theory and Decision*, Jan 2007.

Paul Resnick and Richard Zeckhauser. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In Michael R. Baye, editor, *The Economics of the Internet and E-Commerce*, volume 11 of *Advances in Applied Microeconomics*, pages 127–157. Elsevier Science, Amsterdam, 2002.

Paul Resnick, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara. Reputation systems. *Communications of the ACM*, 43(12):45–48, December 2000.

Paul A Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389, November 1954a.

Paul A. Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389, 1954b.

Reinhard Selten. Axiomatic chracterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–62, June 1998.

Georgios Tziralis and Ilias Tatsiopoulos. Prediction markets: An extended literature review. *Journal of Prediction Markets*, 1(1):75–91, 2007.

Mark Walker. A simple incentive compatible scheme for attaining lindahl allocations. *Econometrica*, Jan 1981. URL `http://ideas.repec.org/a/ecm/emetrp/v49y1981i1p65-71.html`.

Rick Wash and Jeffrey MacKie-Mason. Using a minimum threshold to motivate contributions to social computing. Working paper, 2009.

Wikipedia. Wikipedia:Welcoming committee/Welcome to Wikipedia, 2010. URL `http://en.wikipedia.org/wiki/Wikipedia:Welcoming_committee`. Retrieved on April 2, 2010.

Justin Wolfers and Eric Zitzewitz. Prediction markets. *Journal of Economic Perspectives*, 18(2):107–126, 2004.

Fred Woudenberg. An evaluation of delphi. *Technological Forecasting and Social Change*, 40:131–150, 1991.

H. Peyton Young. Cost allocation, demand revelation, and core implementation. *Mathematical Social Sciences*, 36:213–228, 1998.