Self-Interested Information Providers and Selective Information Disclosure in Multi-Agent Systems

SHANI ALKOBY

Department of Computer Science

Ph.D. Thesis

Submitted to the Senate of Bar-Ilan University

Ramat-Gan, Israel July 2017

This work was carried out under the supervision of Prof. David Sarne, Department of Computer Science, Bar-Ilan University.

Acknowledgments

I would like to express my sincere acknowledgment of the support and help of my advisor Professor David Sarne. His perpetual energy and enthusiasm motivated me greatly during this endeavor. I have learned much from his approach to research and his insightful comments on my work. He was always accessible and willing to help me with my research. As a result, the potential difficulties of research life became smooth and rewarding for me.

Additionally, I would like to thank Sanmay Das, Igal Milchtaich, Esther David, Pingzhong Tang and Zihe Wang for their fruitful cooperation on our joint research projects. I am indebted to my friends in the lab at Bar Ilan University, who helped me enrich my research through long discussions and enjoyable lunches, and to many more colleagues from Israel and abroad who supported me and listened to my ideas even when they were busy.

Last but not least, I would like to thank my beloved family. To my parents Israel and Chani, whose support and encouragement during the past few years was unconditional. To my sisters, Talia and Yaara, and my brother, Ofri, who offered a sympathetic ear to any problem which arose. Their support and encouragement are what made this dissertation possible. My final thanks are to my wonderful husband, Oshri, and my two bright and intelligent little girls, Noga and Shiri, who have made the past few years of intense research and hard work more enjoyable and filled me with happiness and laughter, and by doing so contributed a great deal to the quality of my work.

July 2017 Shani Alkoby

Contents

List of Tables

List of Figures

Abstract

This thesis provides a broad view of the role of information and information's effective providing in multi-agent systems. In the thesis I address three important and highly applicable general settings, exploring how they are influenced by the information provided.

First I discuss the role of information's effective providing in auctions, studying how different information disclosure schemes affect the behavior and expected profit of the different participants in the auction (i.e., information provider, auctioneer and bidders) as well as the social welfare. In this domain I am able to show that there are cases where, by disclosing part of the information for free, the information provider can push the market to better equilibrium, providing her with a higher expected profit. An example for such a case can be an auction of an item which its value is uncertain. By eliminate part of the possible values of the item for free, the information provider can actually increase the need for the exact value causing the other participant to be interested in the information she holds. In addition, I study two complementary manipulations available to the information provider and the auctioneer (such as anonymously disclosing some of the information for free). Using those manipulations, the information provider and the auctioneer can increase their expected profit (and sometimes even the other participants' expected profit). In order to overcome the computational difficulty of extracting the proper manipulation to be used, I provide effective sequencing heuristics that guide the players towards the solution to be evaluated first.

The second domain discussed in this thesis is economic search (in particular one-sided search). In this domain I analyze the case where the information provider is willing to disclose information for free for some chosen outcomes, showing that by doing so she can actually increase her expected profit and sometimes even the searcher's expected profit.

Finally I study the case in which people are in the role of the information buyers, discussing the differences between the results to be achieved when facing completely rational agents and those achieved in experiments performed with people. One interesting result achieved in this domain is that the main reason for those differences is the fact that people fail to take into consideration the strategic nature of their interaction with the information provider and therefore fail to update correctly the posterior probabilities after receiving the free information. People's inability to calculate properly the value of the information was found to be secondary in its influence.

The research summarized in this thesis is based on both theoretical analysis and online empirical experiments. The theoretical analysis is carried out using concepts from game theory, auction theory and search theory, while the online experiments are based on Amazon Mechanical Turk, a well-known crowdsourcing platform.

Chapter 1

Introduction

1.1 Overview

“There will be very few occasions when you are absolutely certain about anything. You will consistently be called upon to make decisions with limited information. That being the case, your goal should not be to eliminate uncertainty. Instead, you must develop the art of being clear in the face of uncertainty."
Andy Stanley *senior pastor of North Point Community Church*

Finding effective ways to provide information is increasingly gaining importance, primarily due to the key role that information is now playing in every aspect of our daily lives. A proof for this can be found in the increasing number of firms investing a big part of their resources in information technologies. For example, in 2014, Facebook bought the social network “Whatsapp" for a sum of $22 billion, making it one of the largest technologies acquisitions to date [42]. In addition, at least 30% of the companies in Forbes’ “World’s Biggest Public Companies" list have a tight connection to providing information [41]. Actually, according to The Economist, the world’s most valuable resource is no longer oil, but data.

In the existing literature much focus has been placed on studying the role information plays in different environments [86, 27, 28, 53, 19, 102, 37, 21, 44, 62, 56, 76, 82, 121, 25, 49, 39, 118]. In addition, recent advances in information processing and communication technologies have given rise to the emergence of strategic information providers in multi-agent settings. These information providers (typically referred to as *information brokers* or *experts*) are capable of disambiguating much of the uncertainty associated with the different alternatives available to agents (e.g., in search-based markets [86, 27], online dating [28], e-commerce [53]).

When dealing with information providing, many questions need to be answered, for example: What is the value of the information? Should one disclose all the information she is holding or only part of it? In what way should the information be displayed? Should the information be given for free or at a cost? Should the disclosure be anonymous? How should one use the information she owns in order to maximize her expected profit from the system? and many more. In this thesis I study ways to effectively provide information to agents in multi-agent systems whenever the information provider is a self-interested agent. The work presented in the following chapters focuses on three highly applicable settings, *Auctions, Economic Search and Interaction with People*, which have been extensively researched in prior literature. For each domain I first present the role of information in the system, moving on to show how information disclosure can influence the different participants in the system (using a game theoretic analysis whenever possible, i.e., when dealing with rational agents) and providing ways to increase the expected profit of different participants using effective disclosure of information. In addition, for the auctions domain, I also provide some interesting manipulations (such as anonymously providing the information or preventing the information holder from providing exact information) available to the different participants that can improve their individual welfare.

1.2 Auctions

Information in Auctions One domain where the choice of the information available to the different players is of great importance is auctions. In particular, whenever the bidders' valuations of the auctioned item depend on some uncertain property of the auctioned item, e.g., its common value, the detail and completeness of the information disclosed is crucial [63, 19]. The information that might be disclosed can influence the identity of the winning bidder, and consequently the expected profit of the auctioneer from the auction, influencing the amount she will be willing to pay to the information provider in exchange for the information. In this context, an external information provider can be of relevance, whenever the auctioneer herself does not have the information necessary to fully disambiguate the uncertainty associated with the item (e.g., does not have the specific expertise or special equipment required for generating the information) or does not want to disclose such information for her own strategic considerations. For example, an individual selling an antique she found in her attic does not necessarily have the expertise needed in order to determine its authenticity and condition. She can, however contact an expert that knows about these things in order to get this information. When dealing with auctions, although it might seem somehow counter-intuitive, I am able to show that by giving away part of the information for free, the information provider's profit (along with the profit of some other participants) can be improved. The underlying idea behind this approach is to change the general belief of the participants in the system regarding the validity of different possible world states, leading them to an equilibrium that is better for the information provider.

Players' Manipulation In many cases, besides selling the information, additional actions can be taken by the player holding the information, allowing her to change the beliefs of other players regarding the true value of the auctioned item. The auctioneer, although not in possession of the information, can also influence the beliefs of the bidders regarding the true value of the auctioned item. In this thesis I discuss two of those additional actions (which I will refer to as manipulations hereafter). The first kind of manipulation is the one carried out by the auctioneer, preventing the information provider from being able to distinguish some of the possible values from the others. In doing so, the information she can receive becomes less accurate (compared to the case of receiving the information without using this manipulation). The second manipulation discussed deals with the information provider disclosing information for free without letting the auctioneer and the bidders know that she is the source of the free information. In doing so, the information provider prevents the strategic response of the other participants (due to the fact that they are not aware of her strategic behavior). In such a case, the dominant strategy of the information provider is to eliminate the set of possible values such that her expected profit from the auction will be maximized. The determination of which values to disqualify is computationally exhausting. In order to provide the information provider with a practical tool in order to decide which values to choose, I present two heuristic methods for sequencing plausible solutions that need to be evaluated, such that those associated with higher profit are sequenced early in the sequence. These heuristics are found to be highly effective experimentally. In addition, I show that initially disclosing the information to all participants instead of just to the auctioneer can sometimes be beneficial to the information provider even though the auctioneer is the only one capable of purchasing the information.

1.3 Economic Search

One of the great successes of the Internet has been in reducing the inherent costs of acquiring information of all kinds. This is particularly true for information platforms of the kinds that connect users with the types of opportunities that they are potentially interested in. These platforms often, either implicitly or explicitly, guide a process of search carried out by users. For example, e-commerce platforms like eBay make it easy to search for consumer goods; Carfax.com makes it easy to search for used cars; Match.com makes it easy to search for romantic partners. The ease with which these Internet-based platforms allow users in locating relevant opportunities has led to a resurgence of research studying the theory and applications of sequential search, with the understanding that the order-of-magnitude reduction in search costs (particularly the cost of time) changes the game and necessitates new methodologies for analyzing these markets [16, 101, 50].

The common scenario in most research dealing with economic search is a set of possible opportunities available to an agent from which she needs to choose only one. Therefore, this agent (mostly known as “searcher") performs a search in order to maximize her expected profit. The searcher is interested in maximizing her overall expected profit from the process (i.e., to maximize x where x is the value of the opportunity found minus the accumulated search cost). Here I address a model of one-sided economic search, e.g., the case in which an individual search for a job or a product. Although the individual is interested in finding the most suitable match, she also needs to take into consideration the cost (e.g., time) that is being wasted during the search.

In one-sided search, information regarding opportunity quality is of great importance to the agent performing the search. In this thesis I demonstrate that allowing the information provider to use partial free information disclosure can push the searcher to continue the search even in cases where she did not intend to do so in the first place (i.e., without having the information disclosed). This can often lead to a higher profit for the information provider (and sometimes even for the searcher (e.g., in the case of a “lemon" opportunity)). Furthermore, I prove a unique equilibrium structure that holds in this case.

One interesting (one may even say natural) phenomenon that may occur in such cases is that additional customers, who were not interested in paying for the information, will now be interested in using the information that is being provided for free. As a result, the expected profit of the information provider may decrease since she will need to produce additional information while not receiving payment for it. This scenario is also discussed here and I am able to show that in this specific model the information provider’s expected profit does not decrease.

1.4 Information and People

In many real life situations, the agents that are interested in the information are humans (i.e., not completely rational agents). Therefore, towards the end of the thesis I investigate the case where the information buyers are people. The results achieved in the case where the information is being sold to people are compared to those achieved in the case where the information buyers are completely rational agents, in order to emphasize the differences between them. Here I am able to show that although when dealing with rational agents the information provider cannot benefit from using partial free information disclosure (a formal proof is provided), this is not the case when it comes to people. When dealing with people, the fact that part of the information was disclosed for free has a positive effect on the information provider’s expected profit. Interestingly, I find that although people are struggling with the calculation of the correct value of the information, and therefor in many times mistakenly choose to purchase the information, this is not the main reason for the difference between the theoretical and actual choice made.

The main reason for the increase in the information provider's expected profit through free information disclosure, when selling information to people, is the fact that people fail to treat the interaction between them and the information provider as strategic.

1.5 Publications

Some of the results appearing in this dissertation were published in the following journal and proceedings of the refereed conferences:

- S. Alkoby, D. Sarne, and I. Milchtaich. Strategic signaling and free information disclosure in auctions. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), pages 319-327, 2017. [6]

- S. Alkoby and D. Sarne. The benefit in free information disclosure when selling information to people. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), pages 985-992, 2017. [3]

- S. Alkoby, D. Sarne, and S. Das. Strategic free information disclosure for search-based information platforms. In Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems(AAMAS-15), pages 635-643, 2015. [4]

- S. Alkoby and D. Sarne. Strategic free information disclosure for a vickrey auction. In International Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis, pages 1-18. Springer, 2015. [2]

- S. Alkoby, D. Sarne, and E. David. Manipulating information providers access to information in auctions. In Technologies and Applications of Artificial Intelligence, pages 14-25. Springer, 2014. [5]

- D. Sarne, S. Alkoby, and E. David. On the choice of obtaining and disclosing the common value in auctions. Artificial Intelligence, 215:24-54, 2014. [102]

1.6 Structure of the thesis

The remainder of the thesis is organized as follows. In the following chapter, I review related work. Chapter 3 introduces the role of information in auction environments. In contrast to previous works, I extend the strategic capability of the information provider so that in addition to setting the price for which she is willing to sell the information, she is now also able to use partial free information disclosure (in the form of sending signals). Next, in Chapter 4, I take a closer look at the role of information in auctions, allowing some of the auction's participants to use some manipulations on the other players. The chapter is divided into two parts where each part allows a different participant to manipulate the rest. The first section discusses the case where the auctioneer manipulates the information provider, preventing her from being able to remove completely the uncertainty associated with the true value of the auctioned item. The second section deals with the manipulation performed both on the auctioneer and the bidders by the information provider. In this section, the information provider also discloses some of the information for free, this time anonymously (i.e., the auctioneer and the bidders do not know she is behind the disclosure of the information). In both sections I provide a game-theoretic analysis, presenting some interesting and counter-intuitive results.

In Chapter 5, I introduce and analyze an effective way to provide information in a one-sided search environment. In this chapter I show that in cases where the information provider/platform is willing to provide some of the premium services for free, she can actually increase her profit from the search.

In Chapter 6, I offer a comparison between a theoretical model in which an information provider is interested in selling information to a rational agent, with the empirical results for the case where people (who are not completely rational players) are in the role of the buyers. In this chapter I show that when it comes to people, the information provider can benefit greatly from using preliminary partial free information disclosure.

I conclude with a discussion and suggested directions for future research in Chapter 7.

Chapter 2

Related Work

The role of information in multi-agent systems and the way it should be provided in those systems is a very relevant and important research topic nowadays. This is due to the fact that information has a hold on almost every situation in our daily lives: from electronic commerce [122, 67] to matching markets [8, 22], from social networks [110] to medical procedures [52, 100], information and its effective usage has the ability to change and improve our lives. As was mentioned above, this thesis discusses the influence information disclosure has in three main settings: Auctions, Economic Search and Interaction with People.

2.1 Information in Auctions

Auctions are an effective means of trading and allocating goods whenever the seller is unsure about buyers' (bidders') exact valuations of the sold item [71, 74]. The advantage of many auction mechanism variants in this context is in the ability to effectively extract the bidders' valuations [70, 91, 30, 106], resulting in the most efficient allocation. Due to its many advantages, this mechanism is commonly used and researched and over the years has evolved to support various settings and applications such as online auctions [64, 77, 57, 32, 106, 104], matching agents in dynamic two-sided markets [20], resource allocation [89, 88, 31] and even for task allocation and joint exploration [45, 75]. In this context great emphasis has been placed on studying bidding strategies [114, 111, 15], the use of software agents to represent humans in auctions [30], combinatorial auctions [112] and the development of auction protocols that are truthful [20, 32, 31, 12] and robust (e.g., against false-name bids in combinatorial auctions [124]). The case where there is some uncertainty associated with the value of the auctioned item is quite common in auction literature. Most commonly it is assumed that the value of the auctioned item is unknown to the bidders at the time of the auction and bidders may only have an estimate or some privately known signal, such as an expert's estimate, that is correlated with the true value [47, 71, 107]. Many of the works using uncertain common value models assumed asymmetry in the knowledge available to the bidders and the auctioneer regarding the auctioned item, typically having sellers who are more informed than bidders [1, 37]. As such, much recent emphasis was placed on the role of information revelation [33, 38, 44, 62] and corresponding computational aspects [37, 21, 34].

2.1.1 Partial Free Information Disclosure

Chapter 3 introduces a model with an augmented information provider's strategy which enables a priori revelation of some of the information for free through the notion of signaling. This adds much complexity, as now both the auctioneer and the bidders need to take into consideration the strategic behavior of the information provider.

Models where agents can disambiguate the uncertainty associated with the opportunities they consider exploiting through the purchase of information have been studied in several other multi-agent domains, e.g., in optimal stopping domains [120, 97, 95, 98, 96, 14]. Here, the main questions studied were how much costly information it makes sense to acquire before making a decision [85, 99], in particular when additional attributes can be revealed at certain costs along the search path [79, 119]. Relaxation of the perfect signals assumption has also been explored in models of two-sided search [28]. Alas, the entities providing the information in such models usually take the form of matchmakers rather than information providers. Those that do consider a self-interested information provider in these domains, e.g., Nahum et al.[86], focused on the way she should set the price for the information she provides and did not consider the option of free information disclosure [56]. The work described in Chapter 5 suggests an information provider that can provide the true value of an opportunity for free, for some of the signals, showing that such a strategy can benefit the information provider. Nevertheless, the source of the achieved improvement in the information provider's profit is completely different than in this case—the free information was shown to push users to become more picky hence it increased the overall search period and consequently the number of times the information provider's service was required. On the contrary, in this case the value derives from the fact that the value of the remaining information held, and consequently the expected profit, increases. Naturally the model and analysis of these two cases are substantially different. Other justifications for free information disclosure mentioned in prior work are increasing user loyalty and attracting potential users [101]. Finally, much of the existing work in auction literature that considers information revelation either assumes that the auctioneer necessarily obtains the information (or initially holds it) or, when an information provider is considered (e.g., [102]), she was either not allowed to be strategic or she was allowed to be strategic but her strategic behavior was limited to price-setting only.

2.1.2 Player's Manipulation

In the currently existing auction literature, great emphasis is placed on the role of information revelation [84, 93, 33, 38, 43, 44, 62, 63]. In particular, several authors have considered the computational aspects of models in which the value of the auctioned item is unknown to the bidders at the time of the auction and bidders may only have an estimate or some privately known signal, such as an expert's estimate, that is correlated with the true value. In such models, the auctioneer needs to decide on the subsets of non-distinguishable values to be disclosed to the bidders [37, 21, 34]. All of these works assume that the auctioneer necessarily obtains the information and that the division into non-distinguishable groups, whenever applicable, is always a priori given to the bidders. Furthermore, not disclosing any information (signal) is not allowed in these works. The problems presented in Chapter 4, on the other hand, do not require that the auctioneer possesses (or purchase) the information in the first place, and they allow non-disclosure of any value even if the information is purchased.

Auctioneer's Manipulation

In contrast to the above presented works, the problem presented in the first section of Chapter 4 allows the auctioneer to use a manipulation on the information provider. The auctioneer, although not holding the information, can choose to limit the level of detail and precision of the information that the information provider will be able to sell. Doing so (for example, by limiting the information provider's access to some of the data required to determine the exact common value), the auctioneer can sometimes increase her expected profit.

Work in other domains that did consider selective information disclosure, e.g., for comparison shopping agents [55] or for sharing data for user modeling [103], is very different in terms of the principles used, since not dealing with an external entity aiming to maximize it expected profit from selling the information, and cannot be applied in this case. All in all, despite the many prior models that consider a subset of the characteristics of the model described in the first section of Chapter 4, to the best of my knowledge, an analysis that addresses all of the different aspects included in this model does not exist in prior literature.

Information Provider's Manipulation

In the second section of Chapter 4 I extend the above described work to include an additional strategic dimension for the information provider, in the sense of anonymously disclosing some of the information for free. Furthermore, unlike prior work, in this section I deal with the computational aspects of extracting the information provider's strategy. Other related work can be found in the study of platforms that bring together different sides of the market (e.g., dating, or eCommerce platforms). Here, there is much work on the impact on selective information disclosure [53], strategic ordering of the disclosed information [54] and having the platform charge only one of the two participating sides [56], and even cases where consumers are in effect paid to use the platform being studied [101]. My work can be viewed in a similar vein, especially in the context of the information provider subsidizing information provisioning, although the intuitions behind my results are quite different and grounded in the transition between different equilibria rather than in the profit of potentially increasing overall participation.

2.2 Information in Economic Search

Much recent work has focused on studying the dynamics associated with information search in distributed multi-agent system environments, where immediate reliable information about the different opportunities available to the agents is not public [56, 86, 119, 27], and emergent behavior in two-sided markets [7, 101, 118, 51]. One of the main questions investigated within this context is how platforms should price their information services, i.e., who pays, and what fees to charge [25, 49, 39, 118]. Chapter 5 is among the first to consider a richer space of strategic choices for platforms, such as the option to partially disclose information for free [92]. To date, work that considers providing information for free has been limited to providing the information completely free to some users. For example, it has been suggested that platforms could charge only one side in a two-sided market while the other group is allowed to use the platform for free [24]. These models are also different from the one presented in Chapter 5 in the motivation for free service provision. Typically, the motivation in these models is intense competition among the players of one group (e.g., directories such as "yellow pages" that are supplied to readers for free) [7] or how platforms can attract elastic consumers and, as a result, obtain higher prices or more participation from the other side [101]. The work presented in Chapter 5 analyzes partial free disclosure of information at the single user level, with the potential benefit that it may induce further consumption of the paid service.

Much recent work has been dedicated to applying search-theoretic principles in novel domains, e.g., in comparison shopping [116, 61]. The assumption in this line of work is that the provider's sole purpose is to serve the user's needs [81]. This assumption leads to the design or modeling of information providers which favor the user (e.g., buyers, in comparison shopping applications) [48, 87]. Existing work where information providers are modeled as self-interested autonomous entities [67, 68] focuses on the use of the information provider for obtaining the signal itself in settings where signals are noiseless (e.g., price quotes) rather than for supplying complementary information [117]. In contrast, the work presented in Chapter 5 deals with an information provider that is interested in maximizing its expected revenue from the process. Finally, there is rich literature on variations of the secretary problem [40], a classic optimal-stopping online problem. Chapter 5's setting is different in that it involves search costs rather than a limited list of possibilities, and the goal is to maximize expected utility rather than the probability of hiring the best candidate (for more on these differences and models that share some features of both types of problems, see Gilbert and Mosteller [46] and Das and Tsitsiklis [29]).

To the best of my knowledge, none of the one-sided search literature in either search theory or multi-agent systems has considered the market dynamics that result in cases where a self-interested information provider can sometimes choose to disclose her information for free.

2.3 Information and People

In human-computer interaction, much effort has been placed on modeling the user's attentional state in order to reason about the cost of (and consequently the benefit in) requesting information from the user or providing her with some information held by the system [123, 59, 60]. While the underlying value of information calculation in these works is similar to the one presented in Chapter 6, the information provider/requester they consider is fully cooperative in the sense that it attempts to maximize the user's expected benefit instead of its benefit from selling the information to the user, as in Chapter 6.

Much work can be found in the multi-agent literature studying strategic information providers that can disambiguate the uncertainty associated with the opportunities available to agents [102, 85, 13]. These, however, primarily deal with the question of information pricing and do not incorporate the option for selectively disclosing some of the information in order to increase the chance for a purchase. Those that do consider the option to use selective information disclosure (in more complex decision settings, where the method can theoretically matter even when taking the strategic aspect of the interaction) [56], or even those that studied the role of information revelation [33, 38, 44], typically assume that information consumers are fully rational agents.

The idea itself of selective information disclosure that affects people's behavior is not new in general and can be found in various other works [113, 108, 11, 54, 92, 9]. It has been justified in prior literature mainly as means for increasing user loyalty, attracting potential users, inducing repeated service requests or influencing the user's behavior [101, 36]. Nevertheless, to the best of my knowledge, an empirical investigation of the benefit in free information disclosure in order to promote information purchase by people has not been carried out to date.

Chapter 3

Providing Information in Auctions

In this chapter[1] I analyze the problem of strategic information disclosure and signaling by information providers in the context of auctions (specifically for second-price auctions). I provide an equilibrium analysis to the case where the information provider can use signaling according to some pre-committed scheme before introducing its regular (costly) information selling offering. The signal provided, publicly discloses (for free) some of the information held by the information provider. Providing the signaling is thus somehow counter intuitive as the information provider ultimately attempts to maximize her gain from selling the information she holds. Still, I show that such signaling capability can be highly beneficial for the information provider and even improve social welfare. Furthermore, the examples provided demonstrate various possible other beneficial behaviors available to the different players as well as to a market designer, such as paying the information provider to leave the system or commit to a specific signaling scheme. Finally, I provide an extension of the underlying model, related to the use of mixed signaling strategies.

3.1 Introduction

Despite their importance in auctions, the study of profit-maximizing information providers in this domain has been limited, to date, to the price setting problem, i.e., the pricing of the information offered for sale (e.g., see [102]). The choice of what information to disclose in auctions

[1]The work reported in this chapter was published in [6].

was studied only in the context of information available to the auctioneer that can possibly be disclosed to the bidders [33, 38, 37, 21, 44, 62]. The analysis of information disclosure by an external self-interested information provider entity, however, calls for a different analysis framework and may reveal much new insights. For example, it has been shown in various domains that information brokers can gain much by limiting their information offers and its accuracy [26] or even offering some of it for free [101, 56].

In this chapter, I introduce a similar approach to the auction domain, focusing in extending an information provider's strategy space to include signaling that aims to selectively disclose, for free, some of the information she holds. For example, before offering to sell the information she own regarding to the worth of the antique, the expert can disclose that the antique's worth can not be lower than a specific value. The signal is disclosed to the auctioneer and bidders prior to making the decision of whether to purchase the information offered for sale. In doing so, the information provider, at times, fully discloses the information she holds and hence the information is not purchased. Yet, this strategy, as demonstrated during this chapter, substantially improves the price the players will be willing to pay for the information in other cases, hence overall the effect on the information provider's profit is positive.

Contributions The main contribution of the chapter is the demonstration that partial free information disclosure may be beneficial for the information provider, despite the counter-intuitiveness of the action. This is demonstrated by a three-party equilibrium analysis for an information-provider-based auction setting with signaling. Furthermore, I show that the benefit of the information provider is not entirely at the expense of the auctioneer. In addition, I show that in various settings players may find it beneficial to pay the information provider to leave the system entirely or switch to a different strategy. Finally, the chapter offers various extensions to the information provider's strategy, such as the use of mixed signaling and restricting her strategy space.

3.2 The Model

I consider a standard second-price sealed-bid (Vickrey) auction setting where bidders' private values depend on some uncertain value X pertaining to (or characterizing) the auctioned item (e.g., the number of people passing by next to an auctioned ad space).[2] The parameter X may obtain any value from a finite set X^*, where the probability it receives a value x is given by $p(x)$ ($\sum_{x \in X^*} p(x) = 1$). X will be called the state of the world.

Each bidder can be of any type T from a finite set T^*, where the probability of a bidder being of a type t is given by $q(t)$ ($\sum_{t \in T^*} q(t) = 1$). The different types are independent. An agent's type t determines its valuation of the auctioned item (i.e., its private value) for each value x that X may obtain, denoted $V_t(x)$. Finally, it is assumed that all players (information provider, auctioneer and bidders) are familiar with the distributions of X and T and the number of bidders taking part in the auction, denoted n, and that each bidder knows her own type. Similar to recent prior work (both in auctions and other domains) I assume that the uncertainty associated with the value of X can be disambiguated by some agent denoted "information provider" [25, 49, 39]. The information provider can sell this information to the auctioneer. I further assume that if such information is purchased by the auctioneer, then she must reveal it to the bidders as well, e.g., as part of fair information disclosure regulations.

Unlike prior work that also used the above underlying model, this model enables the information provider, in addition to setting the price for her information providing service, to send a signal that partially reveals the information she holds. While I do not put any constraint on the signal itself (i.e., it can have any form and its content can either directly relate or have nothing to do with the actual value of X) I assume the signal becomes public domain in the sense that it is revealed both to the auctioneer and bidders. Furthermore, I assume that the information provider must publicly commit to a specific strategy.

Formally, the information provider's strategy, denoted (M, S, C), specifies a set M of possible messages, a function $S : X^* \to M$ specifying the message $S(x)$ that will be sent when the state of the world is x ($X = x$), and a function $C : M \to \mathbb{R}_+$ that specifies the price $C(m) \geq 0$

[2] A specific case is where X represents the common value of the auctioned item [62, 47, 21, 71] and bidders' private values depend to some extent on that common value.

asked for revealing the true state of the world when the message is $m \in M$.

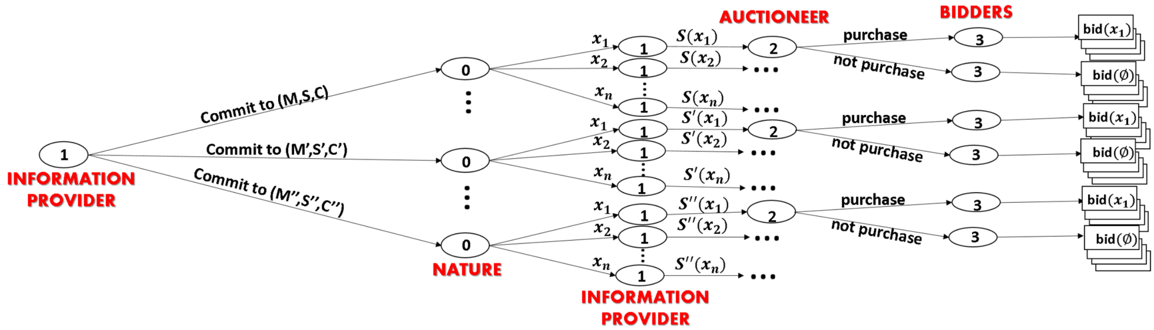


Figure 3.1: Extensive form game representation of the game.

The course of the game is therefore as follows (see Figure 3.1 for the extensive form game representation):

- The Information provider publicly commits to a set of possible signals M, a mapping function S, and a pricing function C. The pair (M, S) will be denoted signaling scheme onward.
- The information provider learns the true state of the world and sends the appropriate signal m according to the signaling scheme she has committed to.
- The auctioneer either purchases from the information provider the information regarding to the true value of X (and truthfully discloses it to the bidders) or does not purchase it.
- Each bidder becomes acquainted with her type and places her bid.

All actions according to the above flow are publicly visible to the other players. Notice that there are several nodes in Figure 3.1 that are in fact in the same information set. For example, it is possible that $S(x') = S(x'')$ (where $x' \neq x''$) hence the nodes of type 3 coming out of the “not-purchase” auctioneer decisions (originating from type 2 nodes) when the information provider commits to a strategy which uses S are all part of the same information set. One important detail that is not being presented in the figure is the fact that the bidders can be of different types hence provide different bids for the same state of the world.

All players are assumed to be fully-rational self-interested agents, aiming to maximize their expected profits. The information provider's profit is her revenue from selling the information. The auctioneer's profit is calculated as her revenue from the auction (captured by the second best bid) minus the payment made to the information provider if the information is purchased. A bidder's profit is the difference between her valuation of the auctioned item and her payment to the auctioneer in case of winning the auction and zero if she loses. Finally, I measure the social welfare as the sum of the auctioneer's, bidders' and the information provider's expected profits. The social welfare is also equal to the expected true valuation of the item in the eyes of the winning bidder. This is due to the fact that both the auctioneer's and the information provider's profits are exclusively based on payments made by or to the other players, thus canceled out by other players' profits, resulting in a social welfare measure that is the true valuation of the item in the eyes of the winner. This represents the efficiency of the allocation made and aligns with the way social welfare is measured in prior work, even when not considering an information provider in the model (Krishna2002 p.75-76).

3.3 Analysis

I analyze the auction using backwards induction. I start with the bidders' best response strategy. A bidder's bidding strategy is influenced by the signaling scheme to which the information provider had committed, the bidder's own type t, the signal m she disclosed or the state of the world x disclosed by the auctioneer. It is captured by the function $B_t : M \cup X^* \to \mathbb{R}$ as follows:

$$B_t(a) = \begin{cases} V_t(a) & a \in X^* \\ \sum_y p(y|a) V_t(y) & a \in M \end{cases} \tag{3.1}$$

where $p(x|m)$ is the conditional probability of $X = x$ given that the signal sent is m, specifically:

$$p(x|m) = \begin{cases} \frac{p(x)}{\sum_{y \in S^{-1}(m)} p(y)} & \text{if } S(x) = m \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

The optimality of the above subscribed strategy derives from the fact that if the information is purchased eventually and the bidders receive the true value x, then this new information necessarily overrides any prior information encapsulated in m. Hence since this is a second price (Vickrey) auction the bidders' best response is necessarily to bid their true valuation, i.e., $V_t(x)$ [115]. Otherwise, if the true value is not purchased by the auctioneer (i.e., $a \in M$) then the bidders should update the probabilities assigned to each possible value $x \in X^*$: (a) each value x for which $S(x) \neq a$ obtains a probability 0 as the information provider's commitment precludes its legitimacy as a potential value X may obtain; (b) the probability of each value x for which $S(x) = a$ is the conditional probability given a, again due to the commitment of the information provider. Based on the updated (posterior) probabilities, the best response strategy is to bid the expected private value [37].

Next, I analyze the auctioneer's strategy. The auctioneer's strategy defines her action to any strategy (M, S, C) used by the information provider and the signal m sent. It needs to take into account the best response strategy of the bidders. I use the function $R_{auc} : M \cup X^* \to \mathbb{R}$ for denoting the expected profit of the auctioneer from the auction (i.e., the second highest bid) when the information provider is committed to (M, S, C) and the bidders use their best response bids. The argument of the function is the true state of the world $x \in X^*$ if the information was purchased and the signal $m \in M$ sent otherwise. The auctioneer's best response is to purchase the information whenever its value is greater than its cost. Formally, the information is purchased whenever $\sum_y p(y|m) \cdot R_{auc}(y) - R_{auc}(m) \geq C(m)$.

Now that the best response strategies of the auctioneer and bidders are defined, I can find an information provider's best response strategy. The information provider will choose a strategy (M, S, C) which maximizes her expected profit, given by:

$$\sum_{x \in X} p(x) \cdot C(S(x)) \tag{3.3}$$

where, for any $m \in M$:[3]

$$C(m) = \max(\sum_y p(y|m) \cdot R_{auc}(y) - R_{auc}(m), 0)$$

[3] Note that for the case where $C(m) = 0$ (i.e., $\sum_y p(y|m) \cdot R_{auc}(y) \leq R_{auc}(m)$, hence the information has no value for the auctioneer) there is an infinite number of best-response strategies for the information provider, as any positive price will lead to the same result of not purchasing the information upon disclosing m.

The sum Equation (3.3) calculates is the expected profit of the information provider when using a strategy (M, S, C) while having every element $C(S(x))$ set to be the maximum possible fee at which the information is purchased by the auctioneer whenever receiving the signal $m \in M$. The calculation sums all possible values in X^* weighing the appropriate gain according to the a priori occurrence probability of each value.

One important feature of the information provider's signaling strategy is that it induces a partition of the set X^*. Two states of the world x_i and x_j, are in the same partition element if and only if the same message is sent in both states. Clearly, the only information revealed by a message is the identity of the partition element that includes the true state of the world; the actual content is irrelevant. Therefore, there is no loss of generality in specifying a strategy as a partition of X^* and a cost c for each partition element.

This observation has two implications. The first is conceptual, as it reveals the main interpretation of the signaling - giving away information. This is further discussed in much detail in the following numerical section. The second implication is computational. Seemingly, the solution concept outlined above would require iterating over an infinite number of signaling schemes. With the observation that the information provider's signaling strategy induces a partition of X^* one needs to consider only a Bell number (of the number of values in X^*) of schemes.[4] This is still intractable when the set of possible values is large, or continuous, however in practice, typically there is a very limited set of world-states (or categories). For example, a geologist selling information about the quantity of oil buried under a land will usually provide you with one out of several ranges. Similarly, the value of a rare coin offered for sale is affected by the era it was made (of a limited set).

The equilibrium can thus be calculated by finding a strategy profile in which all players are using their best response strategy. Since the information provider chooses the solution that maximizes her expected profit and I have already shown that the seemingly infinite strategy space can be reduce to a Bell number, an equilibrium solution necessarily exists.

One key feature of interest in this model where the information provider can use signaling is

[4]The number of possible partitions of a set of size b is a Bell number, given by the recursive formula: $B_{b+1} = \sum_{k=0}^{b} \binom{b}{k} \cdot B_k$, $B_0 = 1$.

the change in the different players' expected profit and in particular the social welfare compared to the case where signaling is precluded. While I discuss and demonstrate numerically typical patterns of changes in the different parties' expected profit in the next section, I can also prove some relationships between the equilibrium social welfare for the two cases.

For this purpose I first define the concept of signaling refinement in the context of signaling schemes in this model, leading to a partial order of equilibria.

Definition 1. *A signaling scheme (M, S) induces a finer partition of the set X^* than the signaling scheme (M', S') if for any x the following holds: $\{y|S(y) = S(x)\} \subseteq \{y|S'(y) = S'(x)\}$ and there exists at least one x for which the inclusion is strict.*

Proposition 1. *Any equilibrium E such that there is no other equilibrium E' that uses a finer signaling scheme is efficient (maximizes the social welfare). In particular, an efficient equilibrium exists.*

Proof. It suffices to show that if there is an equilibrium by which the social welfare is not maximized then there also necessarily exists an equilibrium that uses a finer signaling scheme (hence eventually there is an equilibrium that maximizes social welfare). Consider equilibrium E by which the information provider is using a strategy (M, S, C) and the social welfare is not maximized. Since the social welfare in the auction is equal to the true valuation of the item in the eyes of the winner, the social welfare is maximized whenever the auctioned item is always allocated to the bidder that values it most. In this model this happens whenever all bidders bid their exact valuation according to the true state of the world.[5] Since the social welfare is not maximized in E then there is necessarily a signal $m \in M$ that is used in S for at least two different states of the world ($\exists x_i, x_j \in X^*$ for which $S(x_i) = S(x_j) = m$) and the information is not purchased by the auctioneer upon sending the signal m. Now consider a strategy (M', S', C') which differs from (M, S, C) only in having a different (new) signal for every state of the world for which the signal used with strategy (M, S, C) is m. The expected profit of the information provider is identical with both strategies (M, S, C) and (M', S', C') as the new signals fully disclose the corresponding states of the world and the information is

[5]This can also happen when bidders bid according to expectations however there is one bidder who values the item more than others for each state of the world.

not purchased in those cases. Therefore since (M, S, C) is in equilibrium, so is (M', S', C') (as it maximizes the information provider's profit). The new equilibrium uses a finer signaling scheme than the one used by E by definition. □

One important implication of Proposition 1 is that there exists at least one socially optimal equilibrium. This as opposed to the case of a model where signaling is not used at all, where it is possible that there is no positive value for the auctioneer from the information held by the information provider (hence the information is not purchased). The model where signaling is not used at all is equivalent to the case where the information provider provides an uninformative signal. An uninformative signal is one that encapsulates no information whatsoever, e.g., when always providing the same signal regardless the true world state. Therefore by enabling signaling one can guarantee at least one equilibrium with an improved social welfare, if initially the information offered no value for the auctioneer. Furthermore, the social welfare in this latter case is a lower bound for the social welfare achieved with any signaling scheme as the following proposition states.

Proposition 2. *The social welfare when the bidders do not get any signal or get an uninformative signal is lower than or equal to the social welfare in case of getting any other signal.*

Proof. The information is a random element y. (In the case of the free information (i.e., signaling), $y = S(x)$, where x is the random state of the world). For any fixed vector of bidder types, bidder i's valuation is a random variable $V_{t_i}(x)$ (as it depends on the state of the world). Given the information y, the bidder's bid is the conditional expectation of his valuation, $E(V_{t_i}(x)|y)$. The winning bid is $\max_i E(V_{t_i}(x)|y)$, which is also the conditional expectation of the social welfare, given y. The unconditional expectation is therefore:

$$E(\max_i E(V_{t_i}(x)|y)) \geq \max_i E(E(V_{t_i}(x)|y)) = \max_i E(V_{t_i}(x))$$

The expression on the right-hand is the expected social welfare without the information y. □

On the other hand, if the information is purchased when not using signaling then the social welfare cannot further improve with the use of signaling, as the equilibrium is already efficient.

3.4 Numerical Illustration

I continue by illustrating the benefit for the information provider in free information disclosure (i.e., signaling in this model) and the effect on social welfare and the different players' profit. Since the goal of the numerical examples is primarily illustrative, I use abstract synthetic settings where different bidder types are arbitrarily assigned their private value for any possible state of the world.

n=4

		private values				
		x_1	x_2	x_3	x_4	x_5
Type	p(Values) / q(Types)	0.28	0.19	0.2	0.07	0.26
1	0.38	66	5	35	45	24
2	0.22	72	86	28	73	14
3	0.4	84	14	59	37	81

Table 3.1: The setting used in the example given in Figure 3.2

Consider the auction setting given by Table 3.1. In this example there are four bidders, each assigned type t_1, t_2 or t_3 with probabilities 0.38, 0.22 and 0.4, respectively. The state of the world (the value of X) may obtain one out of five possible values, x_1 through x_5, with the probabilities shown. The remaining values in the table are the private values that bidders of different types assign to the different possible values of X. In this setting, the information provider's expected profit if she decides to commit to the trivial strategy of not disclosing any information through signaling (formally: $S(x_1) = S(x_2) = S(x_3) = S(x_4) = S(x_5)$, or in the shorter form that I will use onwards: $\{\{1,2,3,4,5\}\}$), is 0 since the information is not being purchased by the auctioneer.[6] The information provider can, however, commit to a strategy $S' = \{\{1,3,4,5\},\{2\}\}$, $C(\{1,3,4,5\}) = 1.24$, $C(\{2\}) = 0$, in which case the expected profit is 1.01. This example illustrates the benefit in free information disclosure. The signal results in shrinking the set of possible states of the world, hence the information provider is providing to the other players some of the information she holds, for free. This might seem somehow counter-intuitive, as potentially the information provider could have tried "selling" this information. In particular, whenever disclosing a signal m which is unique, in the sense that

[6] An example where the information is being purchased even when committing to an uninformative signaling scheme is obtained by changing the value of x_1 to bidders of type t_1 in the table from 66 to 200.

there is only one value $x \in X^*$ that maps to it (as in the case of x_2 in the example above), the disclosure of the signal fully reveals the true state of the world and the information provider's service is necessarily not used. Still, by distinguishing this case, the information held by the information provider in other states of the world becomes of greater value for the auctioneer and this added benefit outweighs the loss incurred by giving away part of the information for free. Specifically, in this example, the auctioneer is willing to purchase the information, whenever it is not x_2, for a payment of 1.24. This can be intuitively explained by the fact that bidders of types t_1 and t_3 (the two types associated with a substantial probability compared to t_2) have a relatively low value for x_2. In the absence of indication concerning whether or not $X = x_2$ there is a chance that if purchasing the information the value will turn to be x_2 in which case the bidders of these two types will place low bids, resulting in low expected second best bid. Therefore, while the expected second best bid for all other values will improve, the substantial decrease in profit in case the value x_2 is obtained completely precludes purchase. However, with the initial indication whether x_2 is possible or not, the auctioneer can choose to purchase the information whenever knowing that x_2 is not a possible outcome. Therefore committing to a strategy that gives away some of the information held by the information provider through signaling can be highly beneficial.

I emphasize that in a 2-player setting of an information provider and a potential buyer, where the information offered by the information provider pertains to the true state of the world, giving away free information of this kind cannot be beneficial for the information provider. The proof is similar to the one used in Proposition 2. In the model analyzed in this chapter, however, the free disclosure of information through signaling influences the bids to be placed by the bidders in case the information is not purchased. This directly affects the value of information for the auctioneer and consequently her decision to purchase the information.

Figure 3.2 depicts the players' expected profit and the social welfare, as a function of the strategy used by the information provider for the above setting. The 52 strategies, which is the Bell number for the five values that X may obtain, are aligned along the horizontal axis according to their expected profit to the information provider (ascendingly). The bidders' expected profit in this figure is the sum of the individual expected profits weighted according to

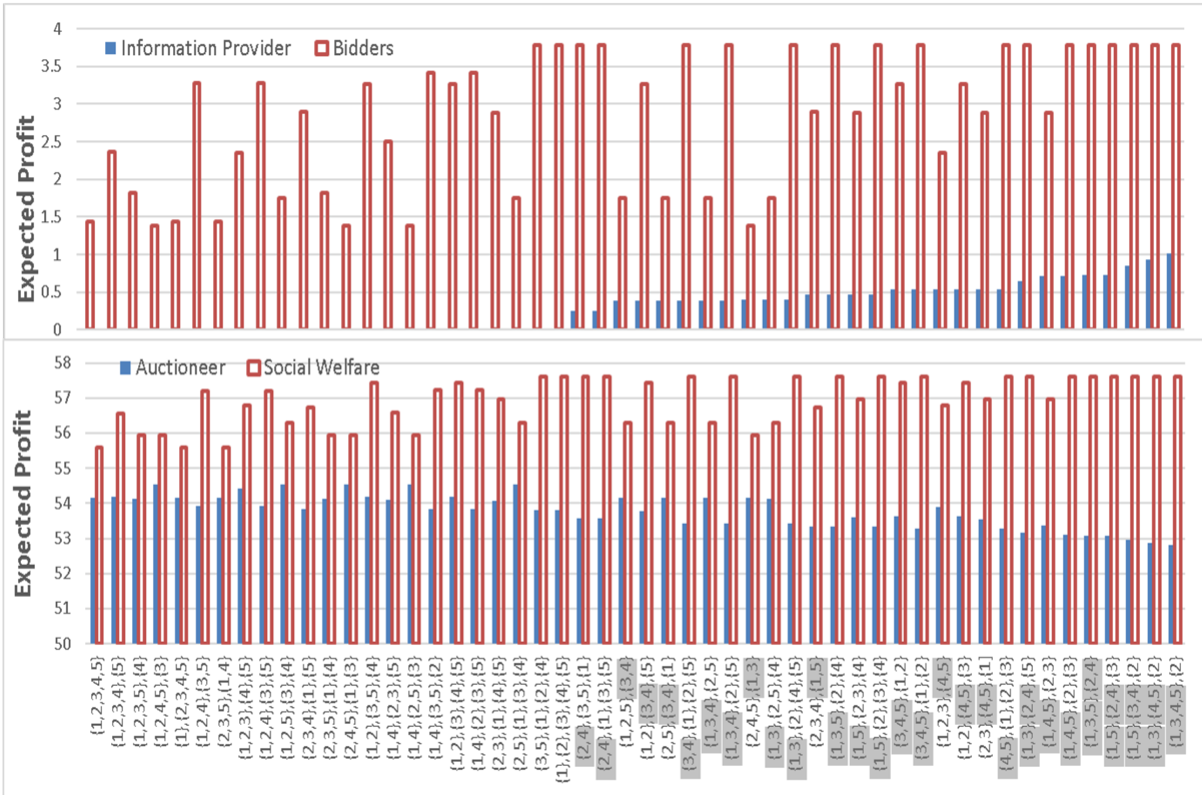


Figure 3.2: The players' expected profit and social welfare for the different signaling schemes the information provider can commit to. The partition elements for which the information is purchased are highlighted.

the types distribution.

As mentioned above, in Figure 3.2 the information is not being purchased when using the strategy $\{\{1, 2, 3, 4, 5\}\}$ and therefore the social welfare associated with this equilibrium is a lower bound to those obtained with any other strategy (see Proposition 2). The social welfare is maximized for all signaling schemes in which the true state of the world is always revealed (i.e., either when it is necessarily purchased (e.g., $\{\{1, 3, 5\},\{2, 4\}\}$), in partitions where it is either purchased or revealed through signaling (e.g., $\{\{3, 4, 5\},\{1\},\{2\}\}$) or when fully revealed through signaling (e.g., $\{\{1\},\{2\},\{3\},\{4\},\{5\}\}$)). In this example, the information provider managed to generate profit through signaling, reaching an equilibrium that is not only efficient but also maximizing the bidders' expected profit. The expected profit of the auctioneer, however, actually decreased in comparison to the case where the signaling is uninformative, and the decrease is greater than the corresponding increase in the information provider's profit when switching to informative signaling. The increase in the information provider's profit does not necessarily need to come fully at the expense of the auctioneer (an example for a case where the information provider's profit is higher than the auctioneer's loss can be obtained by changing the value of x_1 to bidders of type t_2 in the table from 72 to 60). This is best explained by considering the two parts in which information that affects the bidders' bids is being revealed. At the signaling stage, the information provider affects the posterior probabilities of the different values, which reflects on the bids to be placed (and hence bidders and auctioneer's expected profit) if the information is not being purchased. At this stage both the bidders' and the auctioneer's expected profit can increase or decrease. This is best illustrated by the strategies on the horizontal axis, in which the information is not being purchased, each resulting in a different profit to the different players (and, of course, a zero profit to the information provider). At the second stage, where the information can be purchased, the auctioneer's expected profit does not change, as the information provider sets the price such that she takes over whatever additional profit the new information creates for the auctioneer. The signaling scheme set by the information provider therefore controls how much she will be able to charge the auctioneer in the second part, and from the auctioneer's point of view, there is no difference between having the second phase or not.

Based on Figure 3.2 I can extract several benefiting behaviors available to the different players. For example, players can benefit from paying the information provider enough to leave the market completely, or, alternatively, to initially commit to a different signaling scheme which in the absence of proper compensation is not optimal. For example, the equilibrium $S = \{\{1,3,4,5\},\{2\}\}$, $C(\{1,3,4,5\}) = 1.24$, $C(\{2\}) = 0$, yields the auctioneer an expected profit of 52.8 and 1.01 to the information provider. Leaving the market (equivalent to using the strategy $\{\{1,2,3,4,5\}\}$ when the information is not purchased and no information is revealed trough signaling) yields the information provider's profit of zero, however the auctioneer's profit is 54.15. Therefore the auctioneer finds it beneficial to pay the information provider slightly over 1.01 in order to leave. Similarly, the auctioneer finds it beneficial to compensate the information provider for the decrease in her expected gain when switching from the equilibrium strategy $S = \{\{1,3,4,5\},\{2\}\}$, $C(\{1,3,4,5\}) = 1.24$, $C(\{2\}) = 0$, to strategy $S' = \{\{1,2,3\},\{4,5\}\}$, $C(\{1,2,3\}) = 0$, $C(\{4,5\}) = 1.64$, i.e., paying her slightly over 0.47

as the auctioneer's expected profit will increase, from 52.8 to 53.9.

n=4		private values				
		x_1	x_2	x_3	x_4	x_5
Type	p(Values) / q(Types)	0.3	0.3	0.21	0.1	0.09
1	0.18	96	57	46	21	41
2	0.01	69	70	86	76	2
3	0.81	72	5	9	72	14

Table 3.2: An example where the bidders benefit from paying the information provider to commit to a different strategy. For details see text.

Table 3.2 describes a setting where the bidders will benefit from paying the information provider to commit to a particular strategy. In this example the information provider can reach her maximal expected profit, 0.66, using twelve different strategies which among others include the strategy $S = \{\{1,3\},\{2,4\},\{5\}\}$, $C(\{1,3\}) = 5$, $C(\{2,4\}) = 2.125$, $C(\{5\}) = 0$. This strategy results in an expected profit of 3.4 for bidders. With the equilibrium strategy $S' = \{\{2,4\},\{1\},\{3\},\{5\}\}$, $C(\{1,3\}) = 8.65$, $C(\{1\}) = 0$, $C(\{3\}) = 0$, $C(\{5\}) = 0$, on the other hand, bidder's expected profit is 3.46. Therefore, the bidders will find it beneficial to pay the information provider any amount smaller than 0.06 in order to make her choose the latter strategy.

Players can also benefit from constraining the information provider's signaling scheme. Up until now, I assumed the information provider may use any signal. In many cases, however, it is possible that the information provider is limited to (or intentionally chooses (and commits to) limit herself to) a certain subset of possible signals. For example, the information provider may be limited only to signals that partition X^* into two subsets (e.g., providing only signals of the form "greater than w" or "lower than w"). Obviously such a restriction cannot improve the information provider's profit as she uses the expected-profit-maximizing strategy anyhow. A constraint over the information provider's strategy space can, however, be beneficial for the other players, and therefore a market designer may find constraining signaling to specific schemes to be appealing. For example, consider the setting used for Figure 3.2. Here, there are some strategies (e.g., $S = \{\{2,4\},\{3,5\},\{1\}\}$, $C(\{2,4\}) = 0.96$, $C(\{3,5\}) = 0$, $C(\{1\}) = 0$) for which the auctioneer's expected profit increases at the expense of the information provider's expected profit, while bidders' expected profit and the social welfare remain the same (all compared to the equilibrium strategy in the non-restricted scenario). Similar examples where strategy restriction can benefit also the bidders and the social welfare can be produced.

Additional interesting phenomenon is related to the effect of an increase in the number of bidders over the information provider's expected profit. Generally, one would expect the information provider's expected profit to increase as the number of bidders increases. This is because by purchasing the information the auctioneer guarantees that the bidders who value the item most will bid their true valuation. Having more bidders thus should increase the profit for the auctioneer, as it is more likely to have more bidders who assign high values to each state of the world. Since the information provider takes over a substantial portion of the auctioneer's surplus from purchasing the information one would expect the information provider's expected profit to increase as a function of the number of bidders taking part in the auction, as well. The following example, however, illustrates that this is not necessarily the case. Assume there exists a type who will bid high value regardless of the state of the world. In such a case, as the number of bidders rise, so is the probability that there will be two bidder from this type participating in the auction. The auctioneer will thus profit no matter what is the true state of the world and therefore will not be interested in purchasing the information from the information provider.

3.5 Mixed Signaling

The information provider can further improve her profit through the use of mixed signaling strategies. In this case the information provider's strategy specifies a set M of possible messages, a stochastic matrix $A_{|X^*|\times|M|}$, where $A[i,j]$ is the probability that the signal being sent is $m_j \in M$ if the state of the world is $x_i \in X^*$ ($\sum_j A[i,j] = 1$), and a function $C : M \to \mathbb{R}_+$ that specifies the price $C(m) \geq 0$ asked for revealing the true state of the world when the message is m. Unlike the case of committing to a pure strategy, here a strategy does not induce a partition of the set X^* and the information revealed by a message does not necessarily disclose the identity of a subset of X^* that includes the true state of the world. Instead, the message m leads to the posterior probabilities of any of the values in X^*, according to a modification of

Equation (3.2)

$$p(x_i|m_j) = \frac{A[i,j]p(x_i)}{\sum_k p(x_k)A[k,j]} \tag{3.4}$$

I illustrate the potential benefit of using mixed signaling by the setting described by Table 3.3. The maximum expected profit the information provider can achieve through a pure signaling strategy is 1.4 (obtained with the strategy $S = \{\{1,3,4\},\{2,5\}\}$, $C(\{1,3,4\}) = C(\{2,5\}) = 1.4$). Now consider an alternative mixed strategy that uses $M = (m_1, m_2)$, $C(m_1) = C(m_2) = 1.43$ and $A = [(0,1),(1,0),(0,1),(0.1,0.9),(1,0)]$. This means that whenever the state of the world is x_1 or x_3, the information provider uses the signal m_2, whenever it is x_2 or x_5 the information provider uses the signal m_1, and in case the state of the world is x_4, the information provider mixes between the signals m_1 and m_2 with probabilities 0.1 and 0.9, respectively. This strategy improves the expected profit of the information provider to 1.43.

n=4		private values				
		x_1	x_2	x_3	x_4	x_5
Type	p(Values) / q(Types)	0.07	0.14	0.25	0.28	0.26
1	0.24	32	53	9	11	14
2	0.41	68	50	19	50	15
3	0.28	5	85	56	93	70
4	0.07	58	82	88	99	0

Table 3.3: An example where the information provider benefits from using mixed signals. For details see text.

3.6 Conclusions

The analysis provided in this chapter enables demonstrating that by augmenting the information provider's strategy to include signaling she can increase her expected profit. Through the use of signaling the information provider imposes herself on the auctioneer such that the information she holds is actually being purchased even in cases where it cannot be sold otherwise. The importance of this finding is in its non-intuitiveness as the essence of the signaling is free disclosure of some of the information held by the information provider.

The transition to a signaling-based strategy in real-life domains does not require much, given the so many channels available nowadays for disseminating information. For example, a strategic information provider will be able to set up a web-page which includes a reduced set of possible values that the antique can be worth. Since this information is being provided for free to anyone interested, it practical role is identical to the one of a public signal. In fact, it is almost impossible to prevent such a strategic behavior and therefore this should be taken into consideration by the auctioneer and bidders when setting their strategies, making this model a realistic one.

Chapter 4

Players' Manipulations in Auctions

In this chapter[1] I provide a different alternative of using the information in an auction as was presented in the last chapter. Here, I am allowing the information provider and the auctioneer to use some manipulations on the other players, causing them to change the way they react, in order to increase their expected profit. I will first discuss the case where the auctioneer uses a manipulation on the information provider causing her to not be able to remove completely the uncertainty regarding to the true value of the auctioned item. Then I will move on to the case where the information provider uses a manipulation on all the other players in the form of anonymous publicly signals.

4.1 Auctioneer's Manipulation

In this section, I focus on environment settings where the information that may be purchased still involves some uncertainty. The equilibrium analysis is provided with illustrations that highlight some non-intuitive behaviors. In particular, I show that in some cases it is beneficial for the auctioneer to initially limit the level of detail and precision of the information she may purchase. This can be achieved, for example, by limiting the information provider's access to some of the data required to determine the exact common value. This result is non-intuitive especially in light of the fact that the auctioneer is the one who decides whether or not to use the services of the information provider; hence having the option to purchase better information may seem advantageous.

[1]The work reported in this chapter was published in [5, 2]

4.1.1 Introduction

Prior work which include self interested information provider in an auction settings assumed strategic behavior on the auctioneer and the information provider sides. However, the auctioneer's strategy was limited to the choice of the information to be disclosed to the buyers [19, 37]. The information provider, although not being limited to only setting the price of the information provided, was fully certain and captured the exact common value [102].

In this section I extend the model given in chapter 3 to the more realistic case, where the information provider cannot guarantee the identification of the true common value, but rather can offer a more precise estimate of this variable. In particular, I focus on the case in which the information provider can only eliminate some of the possible values and cannot fully distinguish between others. For instance, in the example of the antique found in the attic, it is possible that the information provider will be able to classify the antique's worth as “cheap", “average" and “expensive", where each category spans a wide range of possible values.

To this end, this section's contribution is twofold:

- I augment the three-ply equilibrium analysis (considering the strategic behavior of the information provider, the auctioneer and the bidders) to cases where the information provider can reduce the uncertainty associated with the common value rather than provide its true value.

- I illustrate a beneficial, yet somewhat non-intuitive, strategic behavior of the auctioneer. In particular, this behavior is the auctioneer's choice to intentionally limit the information provider's (e.g., the expert) ability to distinguish between values. This becomes possible when the information provider's ability to provide accurate information depends on inputs received from the auctioneer. In the above mentioned antique found in the attic example, the antique finder can prevent the expert from seeing some relevant documents found with the antique, such that the expert can estimate a specific range of possible worth rather than a certain figure from a wider range of values. The non-intuitiveness of doing this is attributed to the fact that at the end of the day the information provider's information is offered for sale to the auctioneer herself, thus by restricting the information provider's ability to distinguish between values the auctioneer restricts herself by not having the choice of purchasing more accurate information.

This section is structured as follows. In the following subsection I provide a formal presentation of the model. Then, I present an equilibrium analysis and illustrate the potential profit for the auctioneer from influencing the accuracy of the information that can be provided by the information provider. Finally, I conclude with a discussion on the main findings.

4.1.2 The Model

This model considers a similar settings to the model presented in Chapter 3 and includes an auctioneer offering a single item for sale to n bidders using a second-price sealed-bid auction (with random winner selection in case of a tie). The auctioned item is assumed to be characterized by some value X (the "common value"), which is a priori unknown to both the auctioneer and the bidders [58, 47]. The only information publicly available with regard to X is the set of possible values it can obtain, denoted $X^* = \{x_1, ..., x_k\}$, and the probability associated with each value, $Pr(X = x)$ ($\sum_{x \in X^*} Pr(X = x) = 1$). Bidders are assumed to be heterogeneous in the sense that each is associated with a type T that defines her valuation of the auctioned item (i.e., her "private value") for any possible value that X may obtain. Here also I use the function $V_t(x)$ to denote the private value of a bidder of type $T = t$ if the true value of the item is $X = x$. It is assumed that the probability distribution of types, denoted $Pr(T = t)$, is publicly known, however a bidder's specific type is known only to herself.

As in Chapter 3, the model assumes the auctioneer can obtain information related to the value of X from an outside source, denoted "information provider", by paying a fee C that is set by the information provider. Similar to prior models (e.g., [102]), and for the same justifications given there, it is assumed that the option of purchasing the information is available only to the auctioneer, though the bidders are aware of the auctioneer's option to purchase such information. In its most general form, the information provided by the information provider is a subset $X' \subset X^*$, ensuring that one of the values in X' is the true common value. This is usually the case when the information provider cannot distinguish between some of the possible outcomes however can eliminate others. Therefore, the information provider will provide a subset $X' \in D = \{X_1, ..., X_l\}$ where D is the set of possible subsets of X^*, each containing values between which the information provider cannot distinguish, such that $\cup_{X_i \in D} X_i = X^*$ and $X_i \cap X_j = \emptyset, \forall i, j$.

In contrast to the model presented in Chapter 3, here if the information is purchased, the auctioneer, based on the subset obtained, can decide either to disclose the information to the bidders or keep it to herself (hence disclosing $\emptyset$). If she discloses the information, then presumably the information received from the information provider is disclosed as is (i.e., truthfully and symmetrically to all bidders), e.g., if the auctioneer is regulated or has to consider her reputation. Finally, it is assumed that all players (auctioneer, bidders and the information provider) are self-interested, risk-neutral and fully rational agents, and are acquainted with the general setting parameters: the number of bidders in the auction, n, the cost of purchasing the information, C, the possible subsets that may be obtained by the information provider, D, the discrete random variables X and T, their possible values and their discrete probability distributions.

The above model generalizes the one found in [37, 21] in the sense that it requires that the auctioneer decide whether or not to purchase the external information rather than assume that she initially possesses it. Similarly, it generalizes the work in [102] in the sense that it allows the information provider to provide a subset of values rather than the specific true value.

4.1.3 Analysis

The analysis uses the concept of mixed Bayesian Nash Equilibrium. Since the auctioneer needs to decide both whether to purchase the information and if so whether to disclose the information received, I can characterize her strategy using $R^{auc} = (p^a, p^a_1, ..., p^a_l)$ where p^a is the probability she will purchase the information from the information provider and p^a_i ($1 \leq i \leq l$) is the probability she will disclose to the bidders the subset received if that subset is X_i. The dominating bid of a bidder of type t, when subset X' is received (including the case where $X' = \emptyset$, i.e., no information is disclosed), denoted $B(t, X')$, is the expected private value calculated by weighing each private value $V_t(x)$ according to the post-priori probability of x being the true common value given the information X', denoted $Pr(X = x|X')$ [37], i.e.:

$B(t, X') = \sum_{x \in X^*} V_t(x) \cdot Pr(X = x|X')$. If the auctioneer discloses a subset $X' \subset X^* \neq \emptyset$ then $Pr(X = x|X') = \frac{Pr(X=x)}{\sum_{y \in X'} Pr(X=y)}$ for any $x \in X'$ and $Pr(X = x|X') = 0$ otherwise. If no information is disclosed ($X' = \emptyset$) then $Pr(X = x|X' = \emptyset)$ needs to be calculated based on the bidders' belief of whether information was indeed purchased and if so, whether that value is intentionally not disclosed by the auctioneer. Assume the bidders believe that the auctioneer has purchased the information from the information provider[2] with a probability of p and that if indeed purchased then if the information received was the subset X_i then it will be disclosed to the bidders with a probability of p_i. In this case the probability of any value $x \in X_i$ being the true common value is given by:

$$Pr(X{=}x|X'{=}\emptyset) = \frac{Pr(X = x)(p(1 - p_i) + (1 - p))}{(1 - p){+}p\sum_{X_j}(1 - p_j)\sum_{y \in X_j} Pr(X = y)} \tag{4.1}$$

The term in the numerator is the probability that x is indeed the true value however the subset it is in is not disclosed. If indeed x is the true value (i.e., with a probability of $Pr(X = x)$) then it is not disclosed either when the information is not purchased (i.e., with a probability of $(1 - p)$) or when purchased but not disclosed (i.e., with a probability of $p(1 - p_i)$). The term in the denominator is the probability information will not be disclosed. This happens when the information is not purchased (i.e., with a probability $(1 - p)$) or when the information is purchased however the auctioneer does not disclose the subset received (i.e., with a probability of $p \sum(1 - p_j) \sum_{y \in X_j} Pr(X = y)$). Further on in this section I refer to the strategy where information is not disclosed as an empty set. The bidders' strategy, denoted R^{bidder}, can thus be compactly represented as $R^{bidder} = (p^b, p^b_1, ..., p^b_k)$, where p^b is the probability they assign to information purchased and p^b_i is the probability they assign to the event that the information is indeed disclosed if purchased and becomes X_i.

In order to formalize the expected second-best bid if the auctioneer discloses the subset X' I apply the calculation method given in [102] but replace the exact value X with a subset X'. I first define two probability functions. The first is the probability that given that the subset disclosed by the auctioneer is X' , the bid placed by a random bidder equals w, de-

[2]Being rational, all bidders hold the same belief in equilibrium.

noted $g(w,X')$, given by: $g(w,X') = \sum_{B(t,X')=w} Pr(T=t)$. The second is the probability that the bid placed by a random bidder equals w or below, denoted $G(w,X')$, given by: $G(w,X') = \sum_{B(t,X')\leq w} Pr(T=t)$.

The auctioneer's expected profit when disclosing the subset X', denoted $ER_{auc}(X')$, equals the expected second-best bid:

$$ER_{auc}(X') = \sum_{w\in\{B(t,X')|t\in T\}} w(\sum_{k=1}^{n-1} n\binom{n-1}{k}$$
$$(1-G(w,X'))(g(w,X'))^k(G(w,X')-g(w,X'))^{n-k-1} \tag{4.2}$$
$$+\sum_{k=2}^{n}\binom{n}{k}(g(w,X'))^k(G(w,X')-g(w,X'))^{n-k})$$

The calculation iterates over all of the possible second-best bid values, assigning to each its probability of being the second-best bid. As I consider discrete probability functions, it is possible to have two bidders place the same highest bid (in which case it is also the second-best bid). For any given bid value, w, I therefore consider the probability of either: (i) one bidder bidding more than w, $k \in 1,...,(n-1)$ bidders bidding exactly w and all of the other bidders bidding less than w; or (ii) $k \in 2,...,n$ bidders bidding exactly w and all of the others bidding less than w.

Consequently, the auctioneer's expected revenue from the auction itself (i.e., excluding the payment C to the information provider), when the auctioneer uses $R^{auc} = (p^a, p^a_1, ..., p^a_k)$ and the bidders use R^{bidder}, denoted $ER(R^{auc}, R^{bidder})$, is given by:

$$ER(R^{auc}, R^{bidder}) = p^a \sum_{i=1}^{l}\sum_{x\in X_i} Pr(X=x)p^a_i \cdot ER_{auc}(X_i)$$
$$+((1-p^a)+p^a\sum_{i=1}^{l}\sum_{x\in X_i} Pr(X=x)(1-p^a_i))\cdot ER_{auc}(\emptyset) \tag{4.3}$$

where $ER_{auc}(X_i)$ is calculated according to Equation (4.2) (also in the case where $X_i = \emptyset$). Consequently the auctioneer's expected benefit, denoted $EB(R^{auc}, R^{bidder})$, is given by $EB(R^{auc}, R^{bidder}) = ER(R^{auc}, R^{bidder}) - p^a * C$.

A stable solution in terms of the mixed Bayesian Nash Equilibrium in this case is necessarily of the form $R^{auc} = R^{bidder} = R = (p, p_1, ..., p_l)$ (because otherwise, if $R^{auc} = R' \neq R^{bidder}$

then bidders necessarily have an incentive to deviate to $R^{bidder} = R'$), such that: (a) for any $0 < p_i < 1$ (or $0 < p < 1$): $ER_{auc}(\emptyset, R) = ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) = ER_{auc}((1, p_1, ..., p_l), R^{bidder})$); (b) for any $p_i = 0$ (or $p = 0$): $ER_{auc}(\emptyset, R^{bidder}) \geq ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) \geq ER_{auc}((1, p_1, ..., p_l), R^{bidder})$; and (c) for any $p_i = 1$ (or $p = 1$): $ER_{auc}(\emptyset, R^{bidder}) \leq ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) \leq ER_{auc}((1, p_1, ..., p_l), R^{bidder})$. The proof for this derivation is similar to the proof given in [102], with the exception that instead of referring to individual values of X I refer to subsets of values X_i. Therefore one needs to evaluate all the possible solutions of the form $(p, p_1, ..., p_l)$ that may hold (where each probability is either assigned 1, 0 or a value in-between). Each mixed solution of these $2 \cdot 3^k$ combinations (because there is only one solution where $p = 0$ is applicable) should be first solved for the appropriate probabilities according to the above stability conditions. Since the auctioneer is the first mover in this model (deciding on whether or not to purchase information), the equilibrium used is the stable solution for which the auctioneer's expected profit is maximized.

I note that if the information is provided for free $(C = 0)$ then information is necessarily obtained and the resulting equilibrium is equivalent to the one given in [37] for the pure Bayesian Nash Equilibrium case and in [21] for the mixed Bayesian Nash Equilibrium case. Similarly, if $|X_i| = 1 \; \forall i$ is enforced (i.e., the information provider provides the exact value of X) then the resulting equilibrium is the same as the one given in [102].

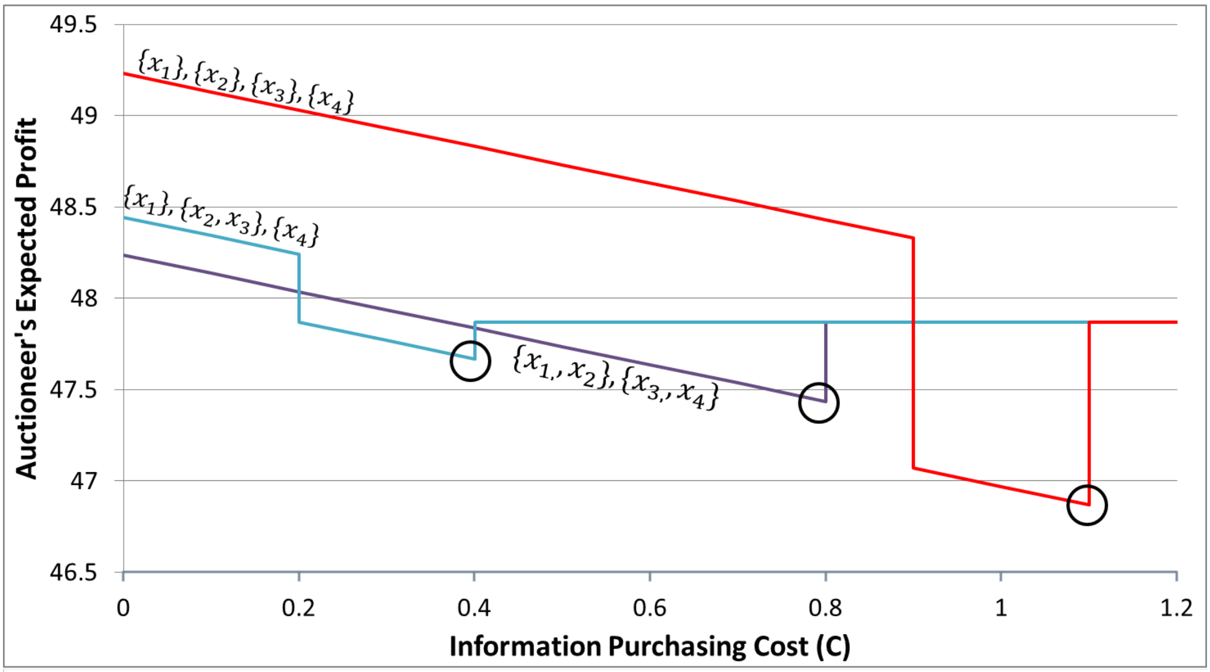
4.1.4 Influencing the Information Provider's Capabilities to Distinguish Between Values

As discussed in the introduction, in various settings the auctioneer can influence the information provider's ability to distinguish between different values the common value obtains. In this subsection I consider the case where the auctioneer has full control over the structure of D, i.e., the division of X^* into disjoint subsets, each composed of values which the information provider cannot distinguish between.

Limiting the information provider's ability to distinguish between values may seem non-intuitive in the sense that it limits the auctioneer's strategy space when it comes to disclosing this information to bidders, if it is purchased. Nevertheless, in many settings the strategy of constraining the information provider's input can actually play into the hands of the auctioneer and improve her expected profit. This happens since when being able to distinguish between values, the information provider will demand more for her services This phenomenon is illustrated in Figure 4.1, which depicts the auctioneer's expected profit (vertical axis) as a function of the information purchasing cost (horizontal axis), for several possible divisions of X^* into subsets of non-distinguishable values. The setting used for this example is given in the table below the graph. It is based on three bidders, where each can be of four different types. The first column of the table depicts the different bidder types and the second column gives their probability. Similarly, the second and third rows depict the different possible values of X (denoted x_1,x_2,x_3 and x_4) and their probabilities. The remaining values are the valuations that bidders of different types assign different possible values of the parameter X. For example, if a bidder is of type 3, then her valuation of x_2 is 59.

Each of the three graphs given in the figure relates to different possible divisions, d of X^* (marked next to it), depicting the expected profit of the auctioneer in the equilibrium resulting in the specific cost of information on the horizontal axis. In this example the resulting equilibrium is always based on pure strategies (i.e., $p, p_i \in \{0, 1\}$) and the points of discontinuity in the curve represent the transition from one equilibrium to another. In particular, for C values in which the curve decreases, the equilibrium is based on always purchasing the information (though not necessarily disclosing all subsets). This happens when the cost of purchasing the information justifies its purchase, i.e., for relatively small C values. The non-decreasing part of the curve is associated with an equilibrium in which the information is essentially not purchased.

As can be seen from the figure, for any cost of purchasing the information $0.9 < C < 1.1$, the auctioneer is better off not allowing the information provider to distinguish between all values: the division $d = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}\}$ is dominated by $d' = \{\{x_1\}, \{x_2, x_3\}, \{x_4\}\}$ and $d'' = \{\{x_1, x_2\}, \{x_3, x_4\}\}$. The explanation for this interesting phenomenon lies in the different costs of the transition between equilibria due to stability considerations. With fully distinguishable values, it is possible that a desired solution which yields the auctioneer a sub-



n=3		private values			
		x1	x2	x3	x4
Type	p(type) \ p(value)	0.626	0.159	0.143	0.072
1	0.252	54	25	84	10
2	0.33	79	56	0	53
3	0.386	21	59	46	69
4	0.032	25	94	69	93

Figure 4.1: The auctioneer's expected profit as a function of the information purchasing cost for different divisions of X^* into subsets of non-distinguishable values.

stantial expected profit is not stable (e.g., in this case when $0.9 < C < 1.1$ the solution is that the information is not purchased at all), whereas with inaccurate information the solution is stable and holds as the equilibrium.

In particular, in this example, when the information provider acts fully strategically, i.e., sets the price of information to the maximum possible price for which the information will still be purchased (the C value in which the equilibrium changes from purchase to not purchase the information, marked with circles in the graphs) the auctioneer will gain (and the information provider will essentially lose) from restricting the information provider's ability to distinguish between values. For example, with $\{\{x_1\},\{x_2,x_3\},\{x_4\}\}$ the information will be priced at $C = 0.4$ yielding the auctioneer an expected profit of 47.6 (compared to $C = 1.1$ and a profit of 46.8 in the "fully distinguishable" case).

4.1.5 Conclusions

In this section I advance the state of the art by providing a three player equilibrium analysis that allows the ability of influencing the auctioneer's expected profit through controlling the granularity and accuracy of the information offered for sale. It is commonly assumed that information providers indeed can control the level of accuracy they offer their customers. Moreover, the accuracy of the information provided depends on the customer's cooperation and the level of the inputs she provides. Against this background, the importance of this equilibrium construction and analysis for auctioneers or the information providers is clear, especially, in terms of the ability to control the granularity in which information is provided.

Here, I show an interesting phenomenon where the auctioneer may benefit in cases where the information provider cannot fully identify the exact state of nature, even though the information is eventually offered exclusively to the auctioneer. This phenomenon is explained by the stability requirement – beneficial solutions that could not hold with the complete ("perfect") information scheme, because of stability considerations, are found to be stable once the information being offered for sale is constrained.

4.2 Information Provider's Manipulation

This section extends prior work by enabling the information broker a richer strategic behavior in the form of anonymously eliminating some of the uncertainty associated with the common value, for free. The analysis of the augmented model enables illustrating two somehow non-intuitive phenomena in such settings: (a) the information broker indeed may benefit from anonymously disclosing for free some of the information she wishes to sell, even though this seemingly reduces the uncertainty her service aims to disambiguate; and (b) the information broker may benefit from publishing the free information to the general public rather than just to the auctioneer, hence preventing the edge from the latter, even if she is the only prospective customer of the service. While the extraction of the information broker's optimal strategy is computationally hard, I propose two heuristics that rely on the variance between the different values, as means for generating potential solutions that are highly efficient. The importance of the results is primarily in providing information brokers with a new paradigm for improving their expected profit in auction settings. The new paradigm is also demonstrated to result, in some cases, in a greater social welfare, hence can be of much interest to market designers as well.

4.2.1 Introduction

Prior work that dealt with uncertain auction settings with a self-interested information broker allowed the information broker to disclose for free part of the information she holds. Doing so, she had to identify herself as the information source. In this section I augment the information provider's strategy, enabling her to anonymously disclose some of the information she holds for free. For example, prior to offering to sell the information she has regarding to the worth of the antique found, the expert can leave an untraceable report eliminating the option of the antique worth being “average". A second somehow surprising choice that I manage to illustrate is the one where the information broker finds it more beneficial to disclose the free information to both the auctioneer and the bidders rather than to the auctioneer only. The latter choice strengthens the auctioneer in the adversarial auctioneer-bidders interaction, allowing her to make a better use of the information offered for sale, if purchased, hence potentially enabling charging more for the service.

As explained in more details in the following paragraphs, the information brokers' problem of deciding what information to disclose for free is computationally extensive. Therefore, another contribution of this section is in presenting and demonstrating the effectiveness of two heuristics for ordering the exponential number of solutions that need to be evaluated, such that those associated with the highest profit will appear first in the ordering.

In the following subsection I provide a formal model presentation. Then, I present an equilibrium analysis for the case where the free information is disclosed to both the auctioneer and the bidders and illustrate the potential profit for the information broker from revealing some information for free, as well as the ordering heuristics and their evaluation. Next, I present the analysis of the case where the free information is disclosed only to the auctioneer. Finally, I conclude with discussion of the main findings.

4.2.2 The Model

This section's model is very similar to the one presented in the previous section, it considers an auctioneer offering a single item for sale to n bidders using a second-price sealed-bid auction (with random winner selection in case of a tie). The auctioned item is assumed to be characterized by some value X (the "common value"), which is a priori unknown to both the auctioneer and the bidders [58, 47]. The only information publicly available with regard to X is the set of possible values it can obtain, denoted $X^* = \{x_1, ..., x_k\}$, and the probability associated with each value, $Pr(X = x)$ ($\sum_{x \in X^*} Pr(X = x) = 1$). Bidders are assumed to be heterogeneous in the sense that each is associated with a type T that defines her valuation of the auctioned item (i.e., her "private value") for any possible value that X may obtain. The use in the function $V_t(x)$ is identical to the one presented in the former section. Again, it is assumed that the probability function of types, denoted $Pr(T = t)$, is publicly known, however a bidder's specific type is known only to herself.

Here also, the model assumes the auctioneer can obtain the value of X from an outer source, denoted "information broker" (for the rest of the section will be called "broker"), by paying a fee C set by the broker. Similar to prior section it is assumed that this option of purchasing the information is available only to the auctioneer, though the bidders are aware of this possibility. In addition, here also, if purchasing the information, the auctioneer can choose whether she is interested to disclose this information to the bidders or keep it to herself (hence disclosing $\emptyset$). Finally it is also assumed that all players (auctioneer, bidders and the broker) are self-interested, risk-neutral and fully rational agents, and acquainted with the general setting parameters.

Up to this point the described model is equivalent to the one found in [102] where the broker is self-interested agent that controls C, the price of purchasing the information. This model, however, extends prior work in the sense that it allows the broker also to anonymously publish some of the information for free before the auctioneer makes her decision of whether to purchase the information. The anonymity requirement in this case is important as discussed later on in the analysis. Yet, there are numerous options nowadays for publishing such information anonymously, e.g., through an anonymous email, uploading the information to an electronic bulletin board or anonymous file server, sending the information to a journalist or an analyst.

The typical case, which I use for the analysis, is the one where the broker, knowing the true value $x \in X^*$, eliminates a subset of values $D \subset X^*$ (where $x \notin D$), leaving only the values $X^* - D$ as applicable values the common value may obtain. Doing so, the model distinguishes between the case where the free information is disclosed to all and the one where it is disclosed to the auctioneer only (allowing the latter to decide what parts of it to disclose further to the bidders prior to starting the auction).

4.2.3 Disclosing Information for Free

Consider the case where the true common value is x. In this case, if the broker publicly eliminates (i.e., anonymously publishes that the common value is not part of) the subset $D \subset X^*$ then the auctioneer and bidders are now facing the problem where the common value may receive only the subset $X^* - D$ and the a priori probability of each value in the new setting is given by $Pr'(X = x) = \frac{Pr(X=x)}{\sum_{x_i \in X^*-D} Pr(X=x_i)}$. Since the auctioneer needs to decide both whether to purchase the true value $x \in X^* - D$ and if so whether to disclose it to the bidders, her (mixed) strategy can be characterized using $R^{auc} = (p^a, p^a_1, ..., p^a_k)$ where p^a is the probability she purchases the information from the broker and p^a_i ($1 \leq i \leq k$) is the probability she discloses to the bidders the value x_i if indeed $X = x_i$. The dominating bid of a bidder of type t, when the auctioneer discloses that the true value is x, denoted $B(t, x)$, is given by $B(t, x) = V_t(x)$ [115]. If no information is disclosed ($x = \emptyset$) then the dominating strategy for each bidder is to bid her expected private value, based on her belief of whether information was indeed purchased and if so, whether the value received is intentionally not disclosed by the auctioneer [37]. The bidders' strategy, denoted R^{bidder}, can thus be compactly represented as $R^{bidder} = (p^b, p^b_1, ..., p^b_k)$, where p^b is the probability they assign to information purchase by the auctioneer and p^b_i is the probability they assign to the event that the information is indeed disclosed if purchased by the auctioneer and turned to be x_i.[3]

The bid placed by a bidder of type t in case the auctioneer does not disclose any value,

[3]Being rational, all bidders hold the same belief in equilibrium.

$B(t, \emptyset)$, is therefore:

$$B(t, \emptyset) = \sum_x V_t(x) \cdot Pr^*(X = x) \tag{4.4}$$

where $Pr^*(X = x)$ is the posterior probability of x_i being the true common value, based on the bidders' belief R^{bidder} and is being calculated as:

$$Pr^*(X = x_i) = \frac{Pr(X = x_i)(p^b(1 - p_i^b) + (1 - p^b))}{(1 - p^b) + p^b \sum(1 - p_i^b)Pr(X = x_i)} \tag{4.5}$$

The term in the numerator is the probability that x_i indeed will be the true value and will not be disclosed. If indeed x_i is the true value (i.e., with a probability of $Pr(X = x_i)$) then it will not be disclosed either if the information is not purchased (i.e., with a probability of $(1 - p^b)$) or if purchased but not disclosed (i.e., with a probability of $p^b(1 - p_i^b)$). The term in the denominator is the overall probability that the information will not be disclosed. This can happen either if the information will not be purchased (i.e., with a probability of $(1 - p^b)$) or when the information will be purchased however the value will not be disclosed (i.e., with probability of $p^b \sum(1 - p_i^b)Pr(X = x_i)$).

Consequently, the auctioneer's expected profit when using R^{auc} while the bidders use R^{bidder}, denoted $EB(R^{auc}, R^{bidder})$, is given by:

$$\begin{aligned} EB(R^{auc}, R^{bidder}) &= p^a \sum Pr'(X = x_i)p_i^a \cdot ER_{auc}(x_i) \\ &+ ((1 - p^a) + p^a \sum(1 - p_i^a)Pr'(X = x_i)) \cdot ER_{auc}(\emptyset) - p^a \cdot C \end{aligned} \tag{4.6}$$

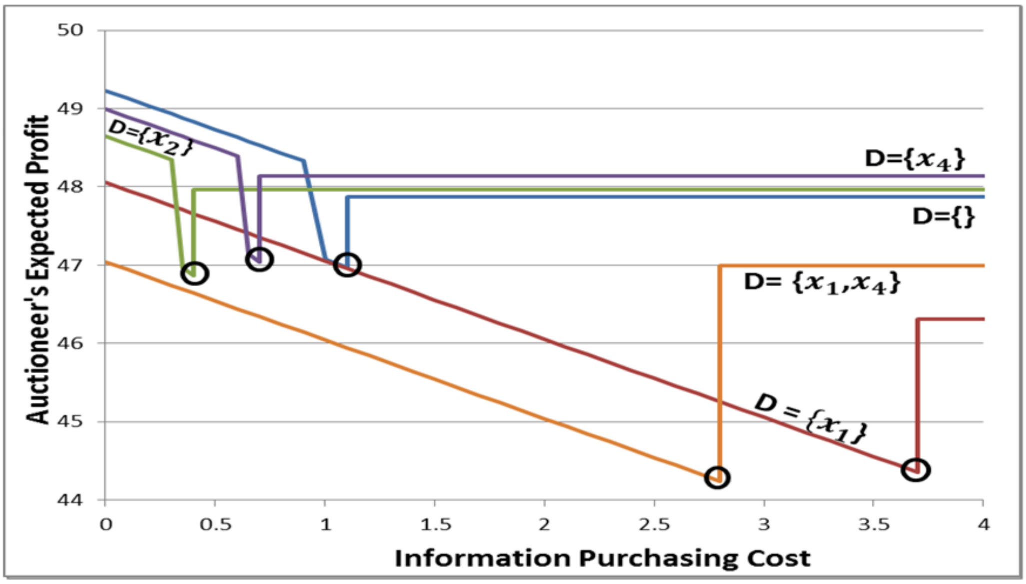
where $ER_{auc}(x_i)$ is the expected second highest bid if disclosing the true value x_i ($x_i \in \{X^* - D, \emptyset\}$). The broker's expected profit is $p^a \cdot C$. The first row of the equation deals with the case where the auctioneer discloses the true value to the bidders (i.e., p^a is the probability that the information was purchased and $\sum Pr'(X = x_i)p_i^a \cdot ER_{auc}(x_i)$ is the probability that x_i is the true value multiplied by the auctioneer's expected profit for this case). The second row deals with the case where the information was not disclosed to the bidders (i.e., when the information is not purchased by the auctioneer (with probability $(1 - p^a)$) and when the information is purchased but not discloses (with probability $p^a \sum(1 - p_i^a)Pr'(X = x_i)$)).

A stable solution in this case (for the exact same proof given in [102]) is necessarily of the form $R^{auc} = R^{bidder} = R = (p, p_1, ..., p_k)$ (as otherwise, if $R^{auc} = R' \neq R^{bidder}$, the bidders necessarily have an incentive to deviate to $R^{bidder} = R'$), such that [102]: (a) for any $0 < p_i < 1$ (or $0 < p < 1$): $ER_{auc}(\emptyset, R) = ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) = ER_{auc}((1, p_1, ..., p_k), R^{bidder})$); (b) for any $p_i = 0$ (or $p = 0$): $ER_{auc}(\emptyset, R^{bidder}) \geq ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) \geq ER_{auc}((1, p_1, ..., p_k), R^{bidder})$; and (c) for any $p_i = 1$ (or $p = 1$): $ER_{auc}(\emptyset, R^{bidder}) \leq ER_{auc}(X_i)$ (or $ER_{auc}(\emptyset, R^{bidder}) \leq ER_{auc}((1, p_1, ..., p_k), R^{bidder})$. Therefore one needs to evaluate all the possible solutions of the form $(p, p_1, ..., p_k)$ that may hold (where each probability is either assigned 1, 0 or a value in-between). Each mixed solution of these $2 \cdot 3^k$ combinations (as only one solution where $p = 0$ is applicable) should be first solved for the appropriate probabilities according to the above stability conditions. Since the auctioneer is the first mover in this model (deciding on information purchase), the equilibrium used is the stable solution for which the auctioneer's expected profit is maximized.

If the information is provided for free $(C = 0)$ then information is necessarily obtained and the resulting equilibrium is equivalent to the one given in [37] for the pure equilibrium case and [21] for the mixed equilibrium case.

Being able to extract the equilibrium for each price C she sets, the broker can now find the price C which maximizes her expected profit. Repeating the process for all different sets $D \subset X^*$, enables extracting the broker's expected-profit maximizing strategy (D, C).

Figure 4.2 depicts the expected profit of the auctioneer (vertical axis) as a function of the information cost C (horizontal axis), for five of the possible D sets. The setting used is given in the table at the bottom of the figure. It is based on four possible values the common value may obtain: $X^* = \{x_1, x_2, x_3, x_4\}$, where x_3 is the true value. The subset D that is used for each curve is marked next to it. For each set D the information provider discloses, the auctioneer chooses whether to purchase the information and what values to disclose, if purchasing, according to the auctioneer's expected-profit-maximizing equilibrium. For example, the lowest curve depicts the auctioneer's expected profit when the broker initially eliminates the values $\{x_1, x_4\}$ and the auctioneer's strategy is to disclose to the bidders the value x_2 in case it is the true value of the auctioned item. Since equilibria in this example are all based on pure strategies, the expected-profit-maximizing price C, and hence the expected profit, equals the highest price at which information is still purchased (marked by circles in the graph, as in this specific example the last segment of each curve applies to an equilibrium by which the information is not being purchased at all). From the figure one can see that indeed in this sample setting, anonymously eliminating some of the applicable values is highly beneficial - for example, the elimination of x_1 results in a profit of 3.7, compared to a profit of 1.2 in the case no information is being a priori eliminated (i.e., $D = \emptyset$).

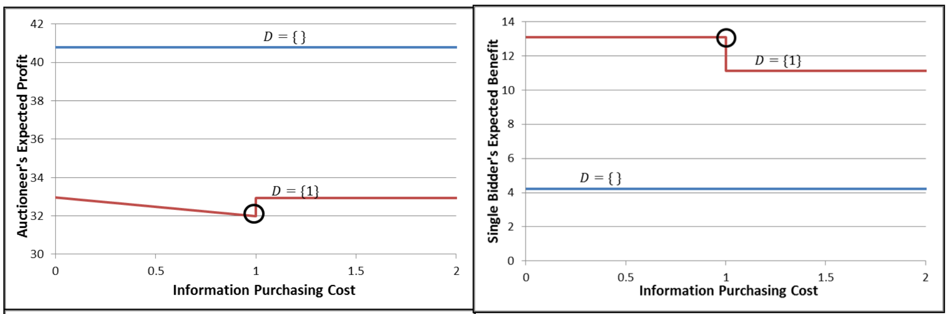


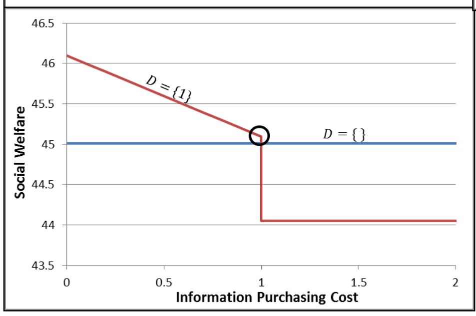
n=3		private values			
		x1	x2	x3	x4
Type	p(type) \ p(value)	0.626	0.159	0.143	0.072
1	0.252	54	25	84	10
2	0.33	79	56	0	53
3	0.386	21	59	46	69
4	0.032	25	94	69	93

Figure 4.2: Auctioneer's expected profit as function of information purchasing cost, for different a priori eliminated subsets.

As discussed in the introduction, benefiting from providing some of the information for free may seem non-intuitive at first—seemingly the broker is giving away some of her ability to disambiguate the auctioneer's and bidders' uncertainty. Yet, since the choice of whether the information is purchased or not at any specific price derives from equilibrium considerations, rather than merely the auctioneer's preference, it is possible that providing information for free becomes a preferable choice for the broker.

The benefit in free information disclosure does not necessarily comes at the expense of social welfare. For exemplifying this I introduce Figure 4.3. The setting used for this example is given in the bottom right side of the figure. Again, the auctioneer's strategy is to disclose the set which will benefit her the most. In this example the broker's expected profit increases from 0 to 1 by publicly eliminating the value x_1 (the information is not purchased otherwise), and at the same time the social welfare (sum of the bidders' and auctioneer's profit) increases from 45 to 45.2, due to the substantial increase in the bidder's profit (from 4.2 to 13.1). If including the broker's expected profit in the social welfare calculation, the increase is even greater.





n = 4

		Private Values			
		x_1	x_2	x_3	x_4
Type	p(value) / p(type)	0.43	0.17	0.3	0.1
1	0.51	43	26	30	42
2	0.25	12	50	71	24
3	0.24	79	7	39	0

Figure 4.3: An example of an improvement both in the broker's expected profit and the social welfare as a result of free information disclosure. The true common value of the auctioned item in this example is x_3.

Finally, I note the importance of disclosing the information anonymously or without leaving a trace of a strategic behavior from the broker's side. If the auctioneer and bidders suspect that the broker may disclose free information strategically, then the equilibrium analysis should be extended to accommodate the probabilistic update resulting from their reasoning of the broker's strategy. This latter analysis is left beyond the scope of the current section—as discussed previously, there are various ways nowadays for anonymous disclosure of information, justifying this specific modeling choice.

4.2.4 Sequencing Heuristics

The extraction of the broker's expected-profit-maximizing subset D is computationally exhausting due to the exponential number of subsets for which equilibria need to be calculated — the broker needs to iterate over all possible $2^{|X^*|-1} - 1$ D subsets (as there are $|X^*| - 1$ values that can be eliminated, and eliminating all but the true value necessarily unfolds the latter as the true one). Therefore, in this subsection I present two efficient heuristics—Variance-based (Vb) and Second-Price-Variance-based ($SPVb$)—that enable the broker to predict with much success what subsets D are likely to result, if eliminated for free, with close to optimal expected profit. The heuristics can be considered sequencing heuristics, as they aim to determine the order according to which the different subsets should be evaluated. The idea is to evaluate early in the process those subsets that are likely to be associated with the greatest expected profit. This way a highly favorable solution will be obtained regardless of how many subsets can be evaluated in total.

Variance-based (Vb)

The value of the information supplied by the broker derives from the different players' (auctioneer and bidders) ability to distinguish the true common value from others, i.e., to better identify the worth of the auctioned item to different bidders. Therefore this heuristic relies on the variance between the possible private values that the information purchased will disambiguate as the primary indicator for its worth. Specifically, if the broker a priori eliminates the subset D, I first update the probabilities of the remaining applicable values, i.e., $Pr^*(x \in X^* - D) = \frac{Pr(X=x)}{\sum_{y \in X^*-D} Pr(X=y)}$. The revised probabilities are then used for calculating the variance of the private values in the bidder's type level, denoted $Var(T = t)$: $Var(T = t) = \sum_{x \in X^*-D} Pr^*(x)(V_t(x) - B(t, \emptyset))^2$, where $V_t(x)$ is the private value of a bidder of type $T = t$ if knowing that the true common value is x, as defined in the model section, and $B(t, \emptyset)$ is calculated according to (Equation (4.4)), based on a setting $X^* - D$. The overall weighted variance is calculated as the weighted sum of the variance in the bidder's type level, using the type probabilities as weights, i.e., $\sum_{t \in T} Pr(T = t) \cdot Var(T = t)$. The order according to which the different subsets $D \subset X^*$ should be evaluated is thus based on the overall weighted variance, descending.

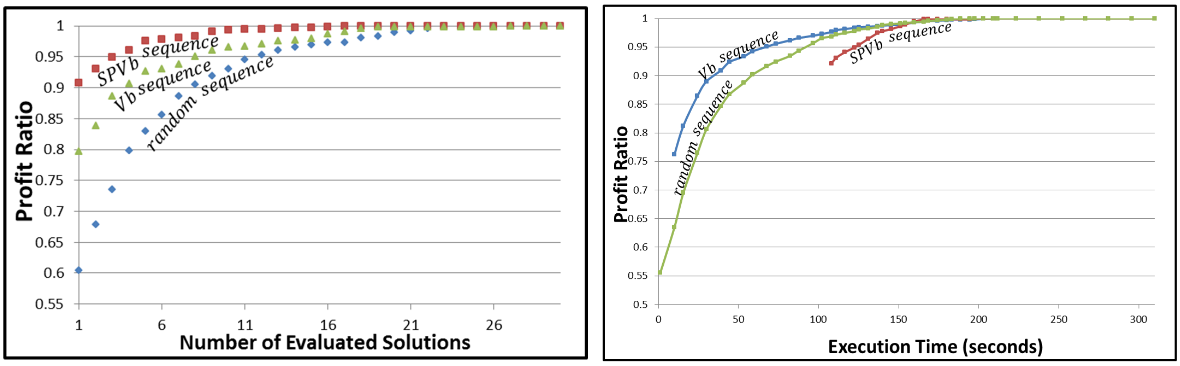


Figure 4.4: Performance (ratio between achieved expected profit and maximal expected profit): (a) Vb and $SPVb$ versus random ordering; and (b) all three methods as a function of running time. All data points are the average over 2500 random settings with 6 possible values the common value obtains.

Figure 4.4(a) illustrates the performance of Vb (middle curve) as a function of the number of evaluated free disclosed subsets (horizontal axis). Since the settings that were used for producing the graph highly varied, as detailed below, I had to use a normalized measure of performance. Therefore I used the ratio between the broker's expected profit if following the sequence generated by the heuristic and the expected profit achieved with the profit-maximizing subset (i.e., how close I manage to get to the result of brute force) as the primary performance measure in the evaluation. The graph depicts also the performance of random ordering as a baseline. The set of problems used for this graph contains 2500 randomly generated settings where the common value may obtain six possible values, each assigned with a random probability, normalized such that all probabilities sum to 1. Similarly, the number of bidders and the number of bidder types in each setting were randomly set within the ranges (2-10) and (2-6), respectively. Finally, the probability assigned to each bidder type was generated in the same manner as with the common value probabilities. For each setting I randomly picked one of the values the common value may obtain, according to the common-value probability function. Each data point in the figure thus represents the average performance over the 2500 randomly generated settings.

As can be seen from the graph, Vb dominates the random sequencing in the sense that it produces substantially better results for any number of subsets being evaluated. In particular, the improvement in performance with the heuristic is most notable for relatively small number of evaluated solutions, which is the primary desirable property for such a sequencing method, as the goal is to identify highly favorable solutions within a limited number of evaluations. As expected, the performance of both Vb and random ordering monotonically increase, converging to 1 (and necessarily reaching 1 once all possible solutions have been evaluated). This is because as the number of evaluated subsets increases the process becomes closer to brute force

Second-Price-Variance-based ($SPVb$)

This heuristic is similar to Vb in the sense that it orders the different subsets according to their weighted variance, descending. It differ from Vb in the sense that instead of depending on the variance in bidders' private values it uses the variance in the worth of information to the auctioneer, i.e., in the expected second price bids. The variance of the expected second price bids if disclosing D for free, denoted $Var(D)$, is calculated as: $Var(D) = \sum_{x \in X^* - D} Pr^*(x)(ER_{auc}(x) - ER_{auc}(\emptyset|D))^2$, where $Pr^*(x)$ is calculated as in Vb, $ER_{auc}(x)$ is the expected second highest bid if disclosing to the bidders that the true value is x, as given earlier. $ER_{auc}(\emptyset|D)$ is the expected second highest bid if the auctioneer discloses no information to the bidders however the bidders are aware of the elimination of the subset D by the broker, i.e., bid according to $B(t, \emptyset) = \sum_{x \in X^* - D} V_t(x)Pr(X = x) / \sum_{x \in X^* - D} Pr(X = x)$.

Figure 4.4(a) also illustrates the performance of $SPVb$ (upper curve) as a function of the number of evaluated subsets D using a similar evaluation methodology and the same 2500 settings that were used for evaluating Vb, as described above. As can be seen from the graph, $SPVb$ dominates random sequencing and produces a substantial improvement, especially when the number of evaluated subsets is small. In fact, comparing the two upper curves in Figure 4.4(a) I observe that $SPVb$ dominates Vb in terms of performance as a function of the number of evaluated sets. One impressive finding related to $SPVb$ is that even if choosing the first subset in the sequence it produces a relatively high performance can be obtained—91% of the maximum possible expected profit, on average. This means that even without evaluating any of the subsets (e.g., in case the broker is incapable of carrying the equilibrium analysis) but merely by extracting the sets ordering, the broker can come up with a relatively effective subset of values to disclose for free.

This dominance of $SPVb$ is explained by the fact that it relies on the variance between the winning bids rather than the bidders' private values. Meaning it relates to the true worth of the information to the auctioneer and consequently to the broker's profit. While this is $SPVb$'s main advantage, compared to Vb, it is also its main weakness: from the computational aspect, the time required for calculating the expected second-price variance of all applicable subsets D is substantially greater than the time required for Vb to calculate the variance between the possible private values. The expected profit of the auctioneer when disclosing the information $X = x$, denoted $ER_{auc}(X = x)$, equals the expected second-best bid when the bidders are given x, formally calculated as:

$$\begin{aligned} &ER_{auc}(X=x) = \sum_{w \in \{B(t,x)|t \in T\}} w(\sum_{k=1}^{n-1} n \binom{n-1}{k} \\ &\sum_{B(t,x)>w} Pr(T=t)(\sum_{B(t,x)=w} Pr(T=t))^k (\sum_{B(t,x)<w} Pr(T=t))^{n-k-1} \\ &+ \sum_{k=2}^{n} \binom{n}{k} (\sum_{B(t,x)=w} Pr(T=t))^k (\sum_{B(t,x)<w} Pr(T=t))^{n-k}) \end{aligned} \tag{4.7}$$

The calculation iterates over all of the possible second-best bid values, assigning for each its probability of being the second-best bid. As I consider discrete probability functions, it is possible to have two bidders placing the same highest bid (in which case it is also the second-best bid). For any given bid value, w, I therefore consider the probability of having either: (i) one bidder bidding more than w, $k \in 1, ..., (n-1)$ bidders bidding exactly w and all of the other bidders bidding less than w; or (ii) $k \in 2, ..., n$ bidders bidding exactly w and all of the others bidding less than w. Notice that Equation (4.7) also holds for the case where $x = \emptyset$ (in which case bidders use $B(t, \emptyset)$ according to Equation (4.4)).

The mentioned calculation results in a combinatorial (in the number of values the common value may obtain) run time. The $SPVb$ method thus requires more time to run for producing the sequence according to which sets need to be evaluated, however the ordering it produces is substantially better than the one produced by Vb. Similarly, random sequencing does not require any "setup" time and the different subsets can be evaluated right away.

In order to weigh in this effect in the heuristics' evaluation I present Figure 4.4(b). Here, the performance is depicted as a function of the actual run-time (in seconds, over the horizontal axis) rather than the number of subsets evaluated once the ordering is completed.[4] Here, one can see the tradeoff between the initial calculation required for the ordering itself and the improvement achieved within the first few evaluated subsets. The shift of each curve over the horizontal axis, till its first data point, is the time it took to generate the sequence of subsets. From the graph one can see that if the amount of time allowed for running is relatively small then one should choose to use a random sequence for evaluation. If the broker is less time-constrained, the best choice is to use Vb and then evaluate subsets according to the generated sequence. One can notice that the same typical behavior was observed for the case of five and seven possible values that the common value may obtain. Evaluating for settings with more than six values is impractical, as it requires solving for thousands of such settings each, as seen from the Table 1, takes substantial time to solve.

Table 1 depicts the average time it took to extract the equilibrium solution for a setting according to the number of values in X^*. Each data point is the average for the 2500 problems described above. This justifies my use of six values settings in the numerical evaluation, and generally motivates the need for the sequencing heuristics I provide by showing that evaluating all possible sets is in many cases impractical — indeed in many cases the total number of values in X^* is moderate,[5] however, even with 8 values it takes more than 10 minutes to extract the broker's equilibrium profit for a single instance.

# of Possible Values	3	4	5	6	7	8
Execution Time (seconds)	0.16	0.58	3.57	20.07	103.19	708.46

Table 4.1: Average time in seconds for extracting the broker's equilibrium profit in a single setting as a function of $|X^*|$.

[4] My evaluation framework was built in Matlab R2011b and run on top of Windows7 on a PC with Intel(R) Xeon(R) CPU E5620 (2 processors) with 24.0 GB RAM.

[5] For example, in oil drilling surveys, geologists usually specify 3-4 possible ranges for the amount of oil or gas that is likely to be found in a given area. Similarly, when requesting an estimate of the amount of traffic next to an advertising space, the answer would usually be in the form of ranges rather than exact numbers.

4.2.5 The Influence of Bidders' Awareness

Next I consider the case where only the auctioneer receives the information disclosed for free (e.g., using anonymous email). In this case the auctioneer needs to decide whether to reveal this information (or part of it) to the bidders. This complicates a bit the structure of the game: (a) First, the broker needs to decide on the set D of values to be eliminated for free and the price C of her service of disambiguating the remaining uncertainty; (b) then, she needs to transfer D anonymously to the auctioneer; (c) next, the auctioneer needs to decide what part $D' \subseteq D$ to further disclose to the bidders; (d) then, the auctioneer needs to decide whether to purchase the true value from the broker, and if purchasing, upon receiving the value, whether to disclose it to the bidders or leave them uncertain concerning the true value; (e) finally, the bidders need to bid for the auctioned item.

The analysis of this case relies heavily on the analysis given above. The resulting adversarial setting if using D and D' is one where bidders bid $V_t(x)$ whenever the information is purchased and disclosed by the auctioneer, and otherwise $B(t, \emptyset)$ according to Equation (4.4), except that this time the probabilities $Pr^*(X = x_i)$ used by bidders result from the equilibrium of a setting where the original values are $X^* - D'$. Therefore, upon receiving the information D from the anonymous source, the auctioneer needs to calculate her expected profit from disclosing any subset $D' \subseteq D$ and choose the one that maximizes it. The auctioneer's expected profit calculation in this case is, however, a bit different, due to the asymmetry in information. When initially disclosing D' to bidders, the auctioneer needs to calculate the expected second best bid from disclosing any value $x \in X^* - D$, based on the bidders' type distribution and their bidding strategy as given above. The auctioneer should choose to disclose any value x for which the expected second best bid if disclosed is greater than the expected second best bid when no information is disclosed (i.e., when bidders bid $B(t, \emptyset)$ according to the equilibrium for the $X^* - D'$ instance of the original problem, as explained above). This allows the broker deciding what subset D to disclose, such that her expected profit is maximized.

Figure 4.5 is an example of a case where the information broker discloses the free information only to the auctioneer and it is to the auctioneer's choice which parts of the information (if at all) to disclose to the bidders prior to the start of the auction. It relies on a setting of three bidders, two possible types and four different values the common value may obtain ($x_1, ..., x_4$), out of which x_4 is the true common value. The full setting details are given in the table in the right hand side of the figure. The leaf nodes provide the expected profit of the auctioneer (inside the rectangle) and the broker (below the rectangle) for each combination of selections made by these two players (the subset D disclosed for free and the subset $D' \subseteq D$ disclosed to the bidders), according to the resulting equilibrium as analyzed above. The yellow colored leafs are therefore those corresponding to the auctioneer's best response given the subset D picked by the broker, hence the expected-profit maximizing strategy for the broker is to anonymously disclose to the auctioneer the subset $\{x_2, x_3\}$ as in this case the auctioneer will choose not to disclose any of these two values to the bidders, resulting in expected profit of 0.9 (compared to 0.8,0.6,0.6,0.8,0.4 and 0.4 if eliminating $\{\emptyset\}$, $\{x_1\}$, $\{x_2\}$, $\{x_3\}$, $\{x_1, x_2\}$ and $\{x_1, x_3\}$, respectively).

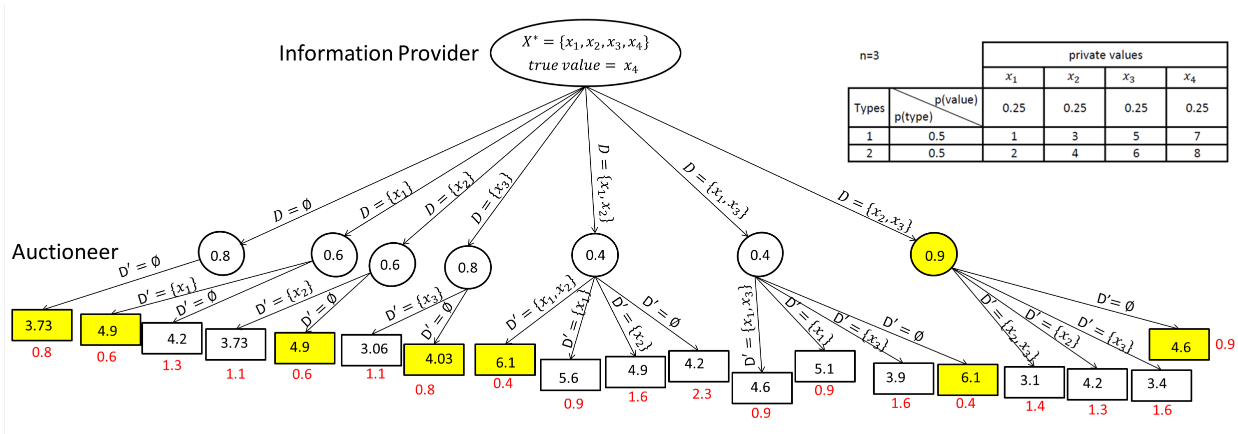


Figure 4.5: Disclosing the free information to the auctioneer only: the broker needs to decide on the subset D to eliminate and then the auctioneer needs to decide on the subset $D' \subset D$ to disclose to the bidders.

Interestingly, if the broker chooses to anonymously disclose to both the auctioneer and the bidders that x_2 and x_3 can be eliminated, her expected profit, calculated based on the analysis given above, is 1.4. This is substantially greater than in the case where the bidders are unaware of the information that was disclosed for free. Furthermore, eliminating x_2 and x_3 for free is not necessarily the broker's expected-profit-maximizing strategy for the scenario where the free information reaches both the auctioneer and bidders. It is possible that there is another subset which elimination results in an even greater improvement in profit when compared to disclosing the elimination of x_2 and x_3 to the auctioneer only. This outcome, as discussed in the introduction is quite non-intuitive because by eliminating the asymmetry in the information disclosed to the different players the broker seemingly reduces the auctioneer's power against the bidders in this adversarial setting. Indeed, when the choice is given to the auctioneer she would rather not disclose this information to the bidders and increase her profit. Since the auctioneer is the potential purchaser of the broker's service information offered by the broker, it might seem that by disclosing the free information only to her, she will have a greater flexibility in making use of the remaining information (that is offered for sale) hence will see a greater value in purchasing it. Yet, the improvement in the auctioneer's competence by disclosing the free information to her only does not translate to an improvement in the broker's profit—eventually the broker's profit depends on the range of prices and the corresponding probabilities at which her information is indeed purchased. These latter factors result from the equilibria considerations, leading to behaviors such as in the example above.

Even for this case, the sequencing heuristics Vb and $SPVb$ are of much importance. Figure 4.6 presents the performance evaluation for these two heuristics, for settings with six values, demonstrating that highly efficient solutions can be extracted even with a small number of evaluations.

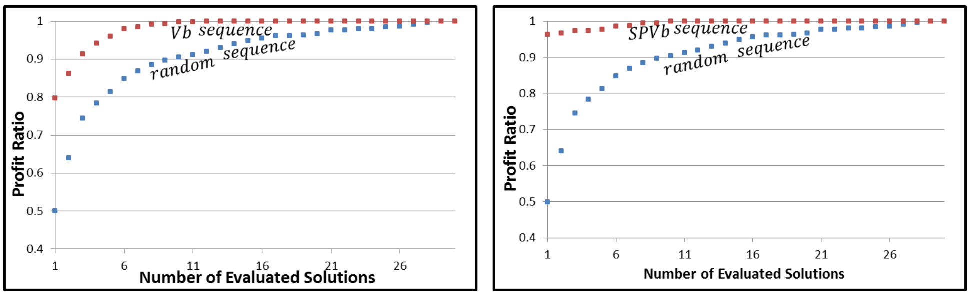


Figure 4.6: Performance (ratio between achieved expected profit and maximal expected profit) when the information is disclosed for free only to the auctioneer and she chooses which information to disclose to the bidders : (a) Vb versus random ordering as a function of number of evaluated subsets; (b) $SPVb$ versus random ordering as a function of number of evaluated subsets. All data points are the average of 2500 random settings with 6 possible values the common value obtains.

4.2.6 Conclusions

The model and analysis given in this section adds an important strategic dimension to prior work in the form of influencing the auctioneer's and bidders' strategic interaction through the anonymous revelation of some of the information that is offered for sale. Throughout the section I show that the use in anonymous disclosure can actually be highly beneficial to the broker. In fact, as demonstrated in the section, it can even lead to an overall improvement in the social welfare. Furthermore, if given the option to disclose the free information to both the bidders and the auctioneer or to the auctioneer only, the broker may benefit from choosing the first, despite the fact that the auctioneer is the one to decide about purchasing the information.

This section presents two sequencing heuristics aiming to reduce the computation time of the broker's expected-profit maximizing strategy. The results of an extensive evaluation of these are quite encouraging - the generated sequences, with both heuristics, are quite effective, as the very few initial subsets placed first in the sequence offer expected profit very close to the expected-profit-maximizing one. Both methods use the variance as a measure for the profit in disclosing a given set, differing in the values based on which the variance is calculated—the bidder's private valuations and the expected second price bids. Interestingly, I find that while the use of the expected second-price produces a substantially more efficient sequence, it is better to rely on the raw values (i.e., bidders' valuations) as the execution time of generating the sequence using the latter method is substantially shorter, leading to better performance overall.

Chapter 5

Providing Information in Search

In this chapter[1] I investigate information platforms that enable and support user search. Consider users engaged in a sequential search process (e.g. for used cars or consumer goods in e-commerce, or partners on a dating website). Many platforms provide basic information on opportunities of interest for free, while also offering, at a price, premium services that can offer more information to the user on the potential values of different opportunities. Prior research has focused on the question of how to price such services. Here I investigate a novel strategic option: can the platform provide some of the premium services for free, and increase its profit in doing so? By analyzing game theoretic equilibria in such a model, I show that there are cases where the platform can indeed benefit by sometimes providing information for free. The underlying mechanism is that sometimes offering free services leads to more extensive usage of the expert's paid services. A robustness analysis shows that even if the population of users is heterogeneous and a large portion of it a priori does not use the premium services, offering parts of the service for free can still be beneficial for the platform despite the potential misuse.

5.1 Introduction

In addition to the development of various information platforms, another concomitant development has been the emergence of a new class of information brokers that serve as intermediaries, typically by helping users to evaluate the relative values of different opportunities that may be

[1]The work reported in this chapter was published in [4]

available to them (for example, Carfax.com in the used-car space, reputation systems in eBay and other auction sites, electronic and human "compatibility consultants" in dating sites).

In many cases, the platform itself offers these information services as part of a "premium package". The typical model of these premium information services is one where users receive noisy signals of the true values of opportunities, and can pay for a premium feature (or external service) that provides more information, helping to disambiguate the uncertainty in the original signal [27, 86].

The study of the strategic behavior of these information intermediaries, whether independent or provided by the platform, has focused primarily on how they should price their services [56, 105, 76, 82, 121, 25, 49, 39, 118]. When intermediaries are paid on a per-use basis (rather than, for example, in commission upon the completion of a transaction), their incentives can become complicated. This is because, for a given user, when the intermediary reveals to the user that an opportunity is a good fit, and the user stops searching and leaves the market, she does not use the intermediary's services any further, cutting off the revenue stream. Therefore, it is typically assumed that the intermediary must be honest for reputation reasons. However, even this, and the literature on this problem thus far, fails to take into account other ways in which the intermediary can remain honest but still increase the probability of extending a user's search process: specifically, it is theoretically possible that the intermediary could sometimes offer to provide extra information for *free* (say for some range of signals received by the user), and, in doing so, actually increase the probability that the user does not terminate her search process and leave the market.

In this chapter, I show that this theoretical possibility is realizable. This chapter's contributions are threefold. First, I provide an equilibrium analysis for a model of sequential search where the platform or external information provider, in addition to choosing the single price it usually charges for its services, can also offer its services for free whenever approached by the searcher. I prove the existence of a unique equilibrium structure in this model and provide the set of equations from which it can be extracted for any given settings. Second, I provide a proof-by-example that free information disclosure can be beneficial. Third, I provide an important robustness-check of the result that free information disclosure can increase profits. The first-order concern, when providing free services, is that a misspecified model of the population can have disastrous consequences – for example, if there exists a group that is characterized by a very low search cost, and members of this group never use the intermediary's services (because the intermediary charges a cost which is too high for them), the intermediary may be unaware of their existence. However, by offering some services for free, the intermediary may expose itself to much higher costs from this group it was previously unaware of. I demonstrate that my example is quite robust to this concern, by showing the percentage of this hidden population would have to be very large to make it unprofitable to use the free information revelation strategy. Taken together, the results suggest that information intermediaries in search-based electronic marketplaces may benefit from disclosing some extra information for free, and that this should be part of the strategic arsenal in algorithmic pricing of information services.

5.2 Model

I consider a standard searcher-platform model (e.g., [56]) in which users, denoted *searchers*, login to the information platform in order to gain access to information about opportunities of the type they seek (e.g., cars, mortgages, consumer products, dates). Due to the high rate of new opportunities arriving to the platform, in practice, one can view it as enabling access to an unlimited stream of opportunities. Each searcher is interested in finding the single best opportunity for them (for example, a searcher would be looking to buy just one used car), so, once they decide on one, I model them as leaving the platform. While unaware of the specific value v of each opportunity listed in the platform, the searcher does know the (stationary) probability distribution function from which opportunities values are drawn, denoted $f_v(x)$. For a cost c_s (monetary, opportunity cost, etc.), the searcher can acquire a signal s, which is correlated with the true value v of an opportunity according to a (known) probability density function $f_s(s|v)$. I assume that higher signals are good news (HSGN), i.e., that if $s_1 > s_2$ then $\forall y, F_v(y|s_1) \leq F_v(y|s_2)$ [83].

The searcher may query and obtain the true value v of an opportunity for which signal s was received, by paying an additional fee c_e. This true value could be obtained from either the information platform that lists the different opportunities or from an external expert (e.g., Carfax.com, or a mechanic). I assume that the platform or expert pays a marginal cost d_e per query (i.e., a "production cost"). For exposition purposes I will use "platform" or "expert" interchangeably to denote this information provider. The goal of the searcher is to maximize the total utility received i.e., the expected value of the opportunity eventually picked minus the expected cost of search and expert fees paid along the way. Thus far, this model is quite standard in prior work [27, 119, 86, 80].

The main departure from previous work in terms of this model is that the expert is allowed to disclose the true value v for free if it determines that this is beneficial. So, for example, if a potential buyer comes to a mechanic with a Carfax report indicating a certain set of flaws, the mechanic may decide to do a free check-up for that car.

I note that the signals received by the searcher are the only form of price discrimination allowed in the model, and thus the only basis on which the free service can be provided in place of the paid service.

Therefore, the model now is as follows. At the very beginning, the expert determines the price it is willing to sell its services for (c_e). Then the search process begins. The searcher receives a signal s; she reveals the signal s to the expert, who must then decide whether to offer its services either for free, or at cost c_e. If it offers the information for free, the searcher takes advantage of the offer, finds out the true value v, and then must decide whether to terminate search and take that opportunity, or to continue search, receiving a new signal s and repeating the process. If the expert chooses not to offer the information for free, the searcher must decide whether to purchase the expert's services at cost c_e. If she does purchase the services, she again finds out the true value v, and then must decide whether to terminate search and take that opportunity, or to continue search. If she does not, then she must decide whether to terminate search and take that opportunity without knowing the true value v, only the signal s, or whether to decline the opportunity and continue search. Figure 5.1 shows the process in the form of a flowchart.

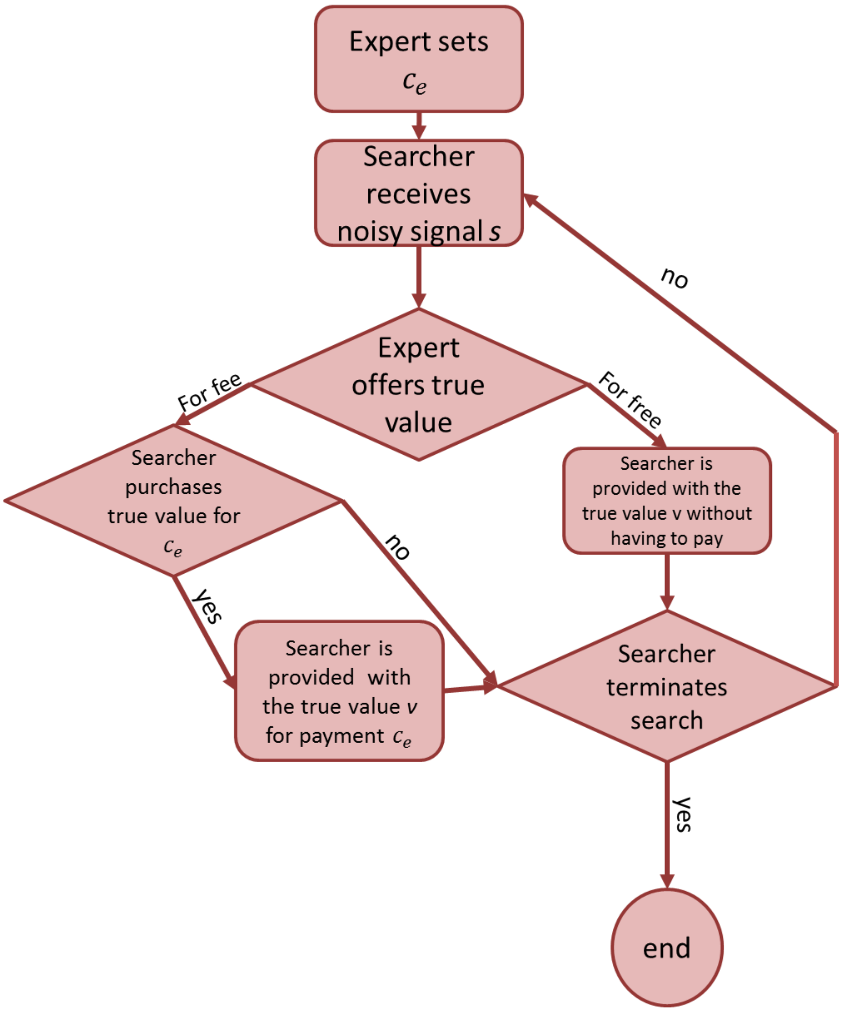


Figure 5.1: Flowchart of the sequential model where the expert may choose to disclose information for free.

5.3 Equilibrium Analysis

No Free Information Disclosure. When the true value is offered by the expert for a fixed fee the game can be solved as a simple Stackelberg game[2] where the expert is the leader, setting the service fee and the searchers are the followers, setting their search strategy accordingly. The searcher in this case, upon evaluating an opportunity and receiving its noisy signal s, can either: (a) reject it and continue search by evaluating a new opportunity; (b) accept it and terminate search; or (c) query the expert to know the true value of the opportunity, incurring a cost c_e, and,

[2] A Stackelberg game is a strategic game in economics in which the *leader* firm moves first and then the *follower* firms move sequentially.

based on the value received, either accept it (terminating the search) or reject it and continue search as before. The optimal strategy for a searcher in this case can be found in prior work (e.g., [80, 27]): it is based on a tuple (t_l, t_u, V) (see Figure 5.2) such that for any signal s: (a) the search should resume if $s \leq t_l$; (b) the opportunity should be accepted if $s \geq t_u$; and (c) the expert should be queried if $t_l \leq s \leq t_u$ and the opportunity accepted (and search terminated) if the value obtained from the expert is above the expected utility of resuming the search, V, otherwise search should resume. This is where V denotes the expected utility-to-go of following the optimal search strategy. The values of t_l, t_u and V can be extracted by solving a set of equations capturing two key indifference situations. The first is where the searcher is indifferent between resuming search and querying the expert (for t_l) and the second when she is indifferent between terminating search and querying the expert (for t_u) [80, 27].

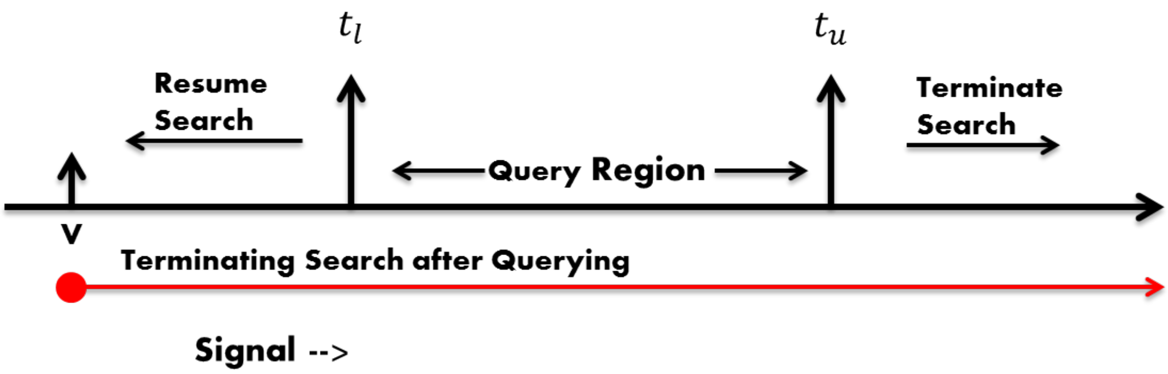


Figure 5.2: Characterization of the optimal strategy for search with an expert (taken from [27]). The searcher queries the expert if $s \in [t_l, t_u]$ and accepts the offer if its value is greater than the value of resuming the search V. The searcher rejects and resumes search if $s < t_l$ and accepts and terminates search if $s > t_u$, both without querying the expert.

With Free Information. When the expert is allowed to offer the true value for some of the signals for free, the equilibrium dynamics become more complex—when setting its service price c_e the expert needs to consider the equilibrium of the simultaneous game resulting from its decision, in which the searcher decides on her search strategy and the expert on the signals for which it will provide the true value for free. The key for solving the problem is therefore understanding the structure of the equilibrium of the resulting simultaneous game given the price c_e set by the expert. Theorem 5.1 provides the structure of the equilibrium for the simultaneous game, showing that it can be compactly represented in the form of four thresholds.

Theorem 5.1. *The equilibrium when the expert is able to disclose information for free, by choice, can be characterized according to the tuple (t_l, t_u, V, t_k) where (see Figure 5.3): (a) the information is offered for free for any signal $t_u \leq s \leq t_k$; (b) the searcher resumes its search for any signal s such that $s \leq t_l$; (c) the searcher accepts any opportunity associated with a signal $s \geq t_k$ and terminate its search right after; (d) the searcher queries the expert for any signal $t_l \leq s \leq t_k$, either for free (if $s > t_u$) or for a cost c_e (otherwise) and accept the opportunity (and terminate search) if the value obtained from the expert is above the expected utility of resuming the search, V, otherwise search is resumed. The values of t_l, t_u, V, t_k can be extracted by solving the set of equations:*

$$V = \frac{-c_s - c_e(F_s(t_u) - F_s(t_l)) + C}{A} \tag{5.1}$$

$$c_e = \int_{y=V}^{\infty} (y - V) f_v(y|t_l) \, \mathrm{d}y \tag{5.2}$$

$$c_e = \int_{y=-\infty}^{V} (V - y) f_v(y|t_u) \, \mathrm{d}y \tag{5.3}$$

$$d_e = \pi_e(F_v(V|t_k)) \tag{5.4}$$

where:

$$A = 1 - F_s(t_l) - \int_{s=t_l}^{t_k} f_s(s) F_v(V|s) \, \mathrm{d}s \tag{5.5}$$

$$\begin{aligned} C = &\int_{s=t_k}^{\infty} f_s(s) E[v|s] \, \mathrm{d}s \\ &+ \int_{s=t_l}^{t_k} f_s(s) \int_{y=V}^{\infty} y f_v(y|s) \, \mathrm{d}y \, \mathrm{d}s \end{aligned} \tag{5.6}$$

$$\pi_e = \frac{(c_e - d_e)(F_s(t_u) - F_s(t_l)) - d_e(F_s(t_k) - F_s(t_u))}{A} \tag{5.7}$$

Proof. I distinguish between three sets of signals. The first, denoted S_{resume}, is the set of signals for which if information is not received for free then the searcher's best response strategy is to resume search without querying the expert. The second, denoted S_{query}, is the set of signals for which even if the information is not free, the searcher's best response strategy is to

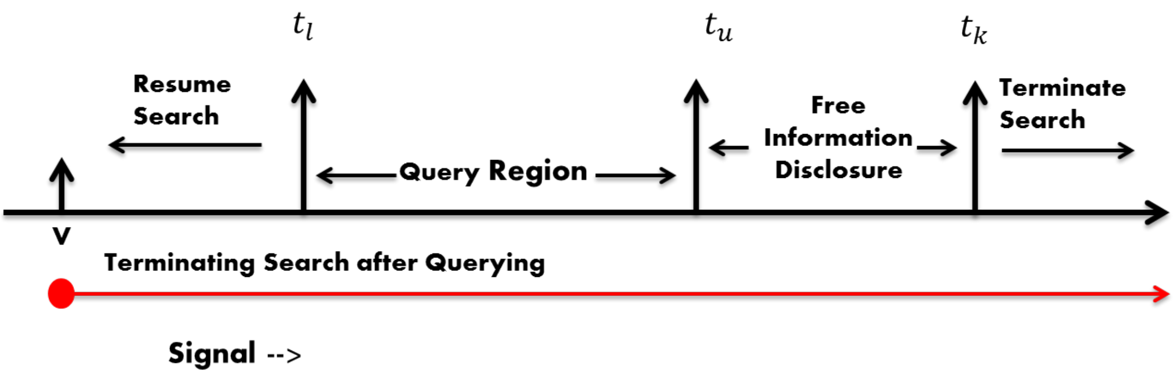


Figure 5.3: Characterization of the optimal strategy for search with an expert when the expert has the option of disclosing part of the information for free. The searcher queries the expert for a fee if $s \in [t_l, t_u]$ and the expert will disclose the opportunity's true value if $s \in [t_u, t_k]$. In both cases the searcher accepts the offer if its value is greater than the value of resuming the search V, and otherwise resumes search. The searcher rejects and resumes search if $s < t_l$ and accepts and terminates search if $s > t_k$, both without querying the expert.

query the expert, and finally the set $S_{terminate}$ denoting the set of signals for which if the information is not free, the searcher's best response is not to query the expert but rather to accept the opportunity and terminate the search. I first prove that from the expert's point of view, if the best response to the searcher's strategy is not to offer the information for free for a signal $s \in S_{terminate}$ then so is the case for any other $s' \in S_{terminate}$ as long as $s' > s$. By providing the information for free when the signal is s the expert incurs a cost d_e, however gains π_e if instead of terminating her search (as is the searcher's strategy for a signal $s \in S_{terminate}$) the searcher, based on the true value received, decides to resume the search. The searcher will decide to resume search only if realizing that the true value is less than the expected benefit of further searching, i.e., if the true value is smaller than V. The probability of the latter event is given by $F_v(v|s)$, hence if the expert prefers not to provide the information for free given signal s then the following must hold:

$$d_e \geq \pi_e(F_v(V|s)) \quad (5.8)$$

Notice that $F_v(v|s) > F_v(v|s')$ for $s' > s$ (due to the HSGN assumption), hence $d_e \geq \pi_e(F_v(v|s)) > \pi_e(F_v(v|s'))$ and therefore the expert necessarily finds it beneficial not to offer the information for free for s'. This is in fact all that needs to be proved for the expert's strategy structure. Obviously there is no benefit from the expert's point of view to offer the information for free for any signal $s' \in S_{query} \cup S_{resume}$ as doing so has no immediate benefit and can only potentially eliminate further search rounds (if the reported true value is greater than V) and future profits.

Moving on to the searcher, I prove that given the above strategy structure of the expert, the searcher's best response strategy is of the (t_l, t_u, V, t_k) structure. First, I prove that given a signal $s \in S_{resume}$, any other signal $s' < s$ also belongs to S_{resume}. The proof is quite straightforward: Let V denote the expected benefit to the searcher if resuming the search if signal s is obtained. Since the optimal strategy given signal s is to resume search, I know $V > E[v|s]$. Given the HSGN assumption, $E[v|s] \geq E[v|s']$ holds for $s' < s$. Therefore, $V > E[v|s']$, proving that the optimal strategy in this case is resuming the search.

Next, I prove that given a signal $s \in S_{terminate}$, any other signal $s' > s$ also belongs to $S_{terminate}$. This proof is also quite straightforward: the searcher decides to terminate the search in case where $E[v|s] > V$. According to the HSGN assumption it is clear that for every $s' > s$ one gets that $E[v|s'] \geq E[v|s] > V$.

The structure of the searcher's strategy, for cases where the information is not offered for free, is thus based on three continuous intervals, represented by (t_l, t_u), where all signals $s < t_l$ belong to S_{resume}, all signals $s > t_u$ belong to $S_{terminate}$ and all signals $t_l < s < t_u$ belong to S_{query}.

At this point, I have everything I need in order to prove that the information will be provided for free only for signals belonging to the continuous interval (t_u, t_k). I have already established the fact that the information provider will never offer the information for free for signals belonging to S_{query} and S_{resume}. Now assume there are signals s and s', such that $s' > s > t_u$ and the expert's best response strategy is to offer the information for free for s' and not for free for s. I have already shown that if $s' > s > t_u$ then both signals belong to $S_{terminate}$. However, if both belong to $S_{terminate}$ and the best response strategy of the information provider is not to provide the information for free for s then, as shown at the beginning of the proof, so is her strategy for s', which leads to a contradiction. Therefore, the set of signals for which information is provided for free is necessarily a continuous interval that starts at t_u.

The searcher therefore will receive the information for free for all signals in the interval

(t_u, t_k) and will query the expert (for a fee) for all signals in the interval (t_l, t_u). In both cases, if the information obtained indicates a value greater than her expected benefit from resuming the search the process will be terminated and otherwise resumed.

Once establishing the general (t_l, t_u, V, t_k) structure I can now formally express the expected profit for the searcher, V, and use optimization for deriving her best response set (t_l, t_u). The searcher's expected profit is given by Equation (5.1). Here the numerator captures the expected profit within a single search round. This is composed by the cost of receiving the signal, c_s, the expected cost of querying the expert, $c_e(F_s(t_u) - F_s(t_l))$, and the expected benefit of the searcher when stopping the search (without taking into consideration the cost of the search or the cost of using the expert), C, as calculated in Equation (5.6). The calculation of C in Equation (5.6) is based on three cases: (i) in case where the value of the signal s is higher than t_k, the searcher's expected profit will be the expectancy of V given the signal ($\int_{s=t_k}^{\infty} \mathrm{f}_s(s)E[V|s]\,\mathrm{d}s$) (ii) in the case where the value of the signal s is in the range of $[t_l, t_u]$ the searcher will stop the search only if the true value of the item is grater than V and in those cases will gain this value ($\int_{s=t_l}^{t_u} \mathrm{f}_s(s) \int_{y=V}^{\infty} \mathrm{y} f_v(y|s)\,\mathrm{d}y\,\mathrm{d}s$) (iii) in the case where the value of the signal s is in the range $[t_u, t_k]$ the searcher again will only stop the search if the item's true value is grater than V and will then gain this true value ($\int_{s=t_u}^{t_k} \mathrm{f}_s(s) \int_{y=V}^{\infty} \mathrm{y} f_v(y|s)\,\mathrm{d}y\,\mathrm{d}s$). I note that since the choice of t_u does not affect C, cases (ii) and (iii) were merge to one integral in Equation (5.3), as will be done in the last two cases of Equation (5.5) to be described. The denominator in Equation (5.1), A, calculated according to Equation (5.5), is the probability that the searcher will terminate the search and purchase the offered item. The searcher will terminate search unless: (i) the value of the signal s is smaller than the value t_l (i.e., with probability $F_s(t_l)$); (ii) the value of the signal s is in the range $[t_l, t_u]$ and the true value of the item is smaller than V (i.e., with probability $\int_{s=t_l}^{t_u} \mathrm{f}_s(s)F_v(V|s)\,\mathrm{d}s$);(iii) the value of the signal is in the range $[t_u, t_k]$ and the true value of the item is smaller than V (i.e. with probability $\int_{s=t_u}^{t_k} \mathrm{f}_s(s)F_v(V|s)\,\mathrm{d}s$).

Setting the first derivative of V according to t_l and t_u to zero obtains Equations (5.2) and (5.3). Finally, Equation (5.4) represents the best response strategy for the auctioneer as explained above.

To conclude the proof I note that there are ultimately 4 strategy parameters: t_l and t_k for the searcher, and t_u and c_e for the expert. Equation (5.2) gives t_l, Equation (5.3) gives t_u, Equation (5.4) gives t_k, and c_e is found by optimizing the expert's profit. □

I note that the above theorem and its proof can be trivially extended for the case where the expert provides a noisy (yet more accurate) signal rather than the true value of the opportunity, using a transformation proposed by MacQueen for the case without the free information disclosure option [80].

Equations (5.2)-(5.4) that characterize the searcher's and the expert's optimal thresholds, can also be derived from their indifference conditions at signals t_l, t_u, and t_k respectively. For example, t_l is the signal at which a searcher is indifferent between either resuming the search or querying the expert, i.e., $V = \int_{y=V}^{\infty} y f_v(y|t_l)\, dy + V F_v(V|t_l) - c_e$, which transforms into Equation (5.2); alternatively, t_l can also be interpreted as a point where cost of purchasing the expert's service is equal to the expected increase in utility from consulting the expert when the searcher would otherwise reject and resume search. Similarly, t_u is the signal at which the searcher is indifferent between querying the expert and terminating the search without querying the expert (in case the information is offered for a fee c_e). Finally, t_k is the signal for which the expert is indifferent between providing the information for free and having the searcher terminate its search, i.e., $0 = -d_e + \pi_e(F_v(v|t_k))$, which transforms into Equation (5.4).

Using the set of Equations (5.1)-(5.7) I can now solve for (t_l, t_u, V, t_k), and in particular Equation (5.7) provides me with the resulting expected profit for the platform. Therefore, the expert can solve for the expected-profit-maximizing c_e (e.g., numerically).

5.4 Numerical Illustration

I can use the characterization of the equilibrium strategies to solve for the expert's optimal service fee and derive implications for how experts should price their services. Equilibrium in expert-mediated search derives from a complex set of dynamics. Many parameters affect the equilibrium, including the distribution of values, the correlation between signals and values, search frictions and the cost of querying the expert. Uncovering phenomenological properties of the model is therefore difficult and restricted using a static analysis. Instead, I turn to an illustrative model that uses a particular, plausible distribution of signals and values. For this purpose I adopt the setting used in Chhabra et al [27]. The setting uses the signal as an upper bound on the true value. So the signal could be thought of as the searcher's optimistic estimate upon observing the opportunity (e.g., sellers and dealers offering cars for sale usually make cosmetic improvements to the cars in question, and proceed to advertise them in the most appealing manner possible, hiding defects using temporary fixes; mortgage lenders may advertise their most appealing features, such as a low introductory rate, while keeping troublesome terms and conditions hidden). Specifically, the signals s are uniformly distributed on $[0, 1]$, and the conditional density of true values is linear on $[0, s]$. Thus

$$f_s(s) = \begin{cases} 1 & \text{for } 0 \leq s \leq 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_v(y|s) = \begin{cases} \frac{2y}{s^2} & \text{for } 0 \leq y \leq s \\ 0 & \text{otherwise} \end{cases}$$

Figure 5.4 depicts the expert's expected profit with and without free information disclosure as a function of the service fee it sets, c_e. The setting used for the graph takes the searcher's search cost to be $c_s = 0.17$ and the expert's production cost $d_e = 0.00019$. Obviously, when $c_e = 0$ the expert makes no profit regardless of whether or not it offers some of the information for free. However as c_e increases, and in particular when $c_e > d_e$ the expert makes profit and, as can be observed from the graph, the option to provide information for free results in a greater expected profit. For larger c_e values ($c_e > 0.028$) the expert becomes too costly and is not being used anymore, i.e., the equilibrium is characterized by $t_l = t_u$.

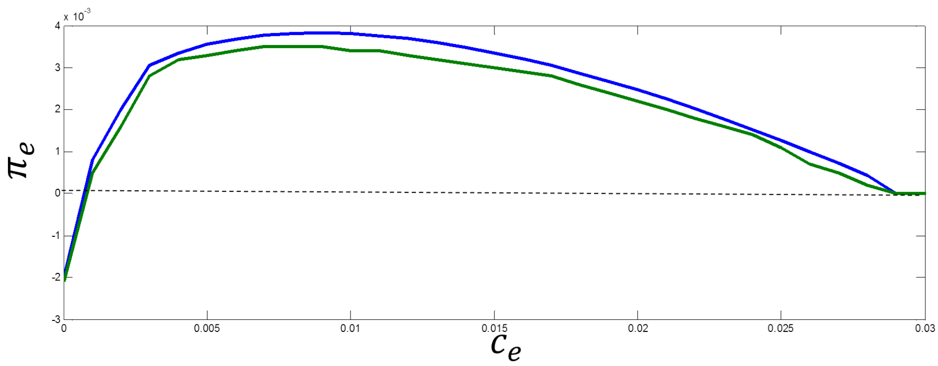


Figure 5.4: Expert's expected profit with and without free information disclosure (upper and lower curve, respectively) as a function of c_e for a setting where: $c_s = 0.17$ and $d_e = 1.9 \cdot 10^{-4}$.

To get a better understanding of the equilibrium dynamics in the resulting simultaneous game once the expert has set its service fee c_e, in particular the effect of free information disclosure on the equilibrium, I present Figure 5.5. The figure depicts the searcher's expected benefit V, the thresholds t_l and t_u and the difference between the two, as a function of the percentage of the interval of signals $(t_u, 1)$ for which information is offered for free, denoted w (i.e., $w = (t_k - t_u)/(1 - t_u)$). I show w on the horizontal axis rather than t_k because an increase in t_k per-se has no actual meaning, as it does not say anything about the higher threshold nor the range of signals used by the searcher for using the costly service (t_u). These result from the equilibrium dynamics of the simultaneous game. The use of w as defined above resolves the problem and enforces an equilibrium in which t_k is constrained in terms of a portion of the resulting $(t_u, 1)$ interval. One possible interpretation for w is therefore the extent to which the expert is willing to provide free information in cases where the searcher receives a favorable signal for which the benefit from knowing the true value does not justify paying c_e for it. The setting used for this figure is the same as the one used for Figure 5.4 ($c_s = 0.17, d_e = 0.00019$), except that here I also fix $c_e = 0.01$, i.e., the expert is not attempting to maximize profits over c_e in that specific market.[3] Figure 5.5 is complemented by Figure 5.6, which depicts the expected number of searches (i.e., expected number of opportunities for which a signal was received by the searcher) and the expected number of times the expert was queried by the searcher in the costly-service mode.

As can be seen from Figure 5.5, the increase in w results in an increase in the searcher's expected profit. This is expected, as the searcher now receives the true value for free for some of the signals and therefore, since c_e has not changed, it cannot possibly do worse than in the case where the information is always costly. The increase in V results in an increase in t_l and t_u as the searcher will now become indifferent to querying the expert for greater signals. While the

[3]This is often the case whenever the expert is operating in parallel markets and needs to set a fixed fee, or cannot distinguish users coming from this market from others.

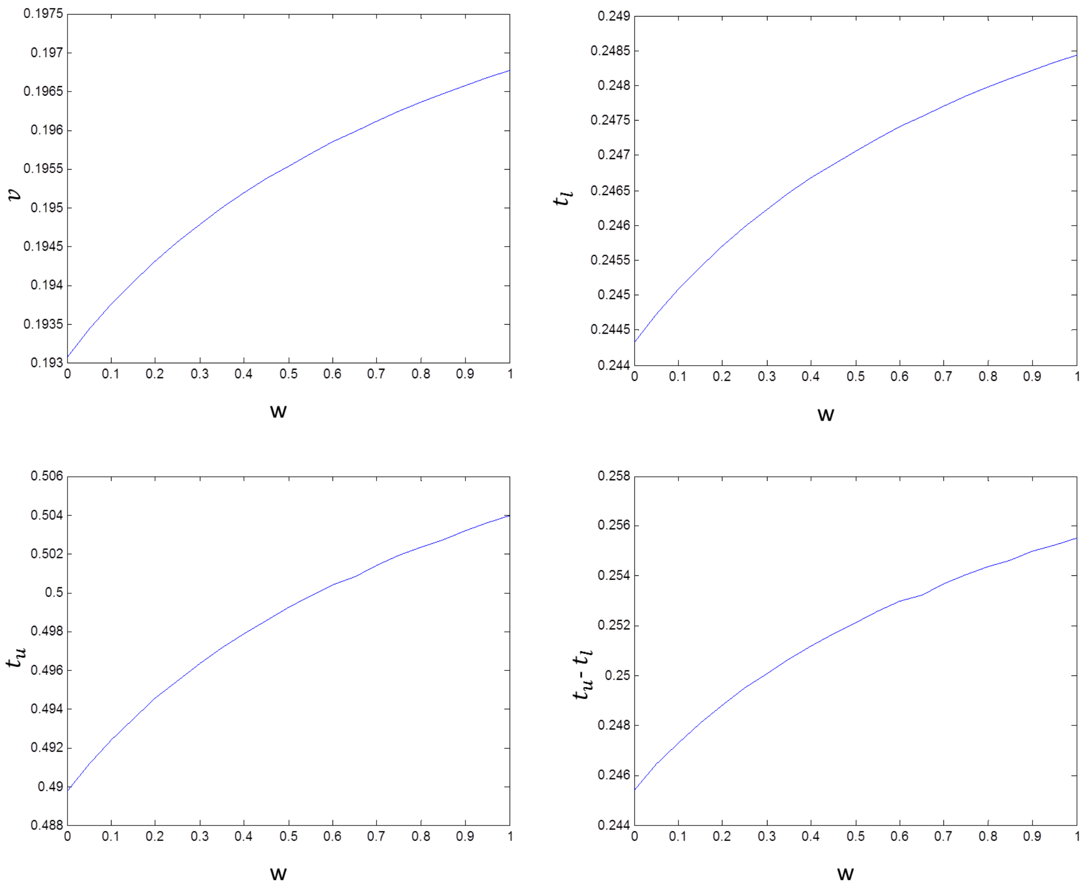


Figure 5.5: V,t_l,t_u and the size of the interval between t_l and t_u as a function of the parameter w (the percentage of the interval of signals $(t_u, 1)$ for which information is provided for free). The setting is $c_s = 0.17$, $d_e = 0.00019$ and $c_e = 0.01$.

increase in t_u is beneficial, from the expert's point of view, as it increases the interval of signals for which the service is used for a payment, the increase in t_l has the exact opposite effect. Fortunately, since in this example I use a uniform distribution of signals, I can rely on the measure $t_u - t_l$ to determine whether or not the probability the expert will be queried for a fee increased. From the figure one can see that indeed the increase in w results, in this example, in an increase in $t_u - t_l$ and consequently an increase in the chance the expert is used for a payment c_e in each search round. Overall, one can see from Figure 5.6 that the increase in w results both in an increase in the expected number of search rounds and in the expected number of queries made for a fee. The increase in the first measure suggests that the searcher has become more picky. This is interesting especially since with the increase in $t_u - t_l$ and the increase in the portion of $1 - t_u$ for which free information is received the searcher receives/purchases more

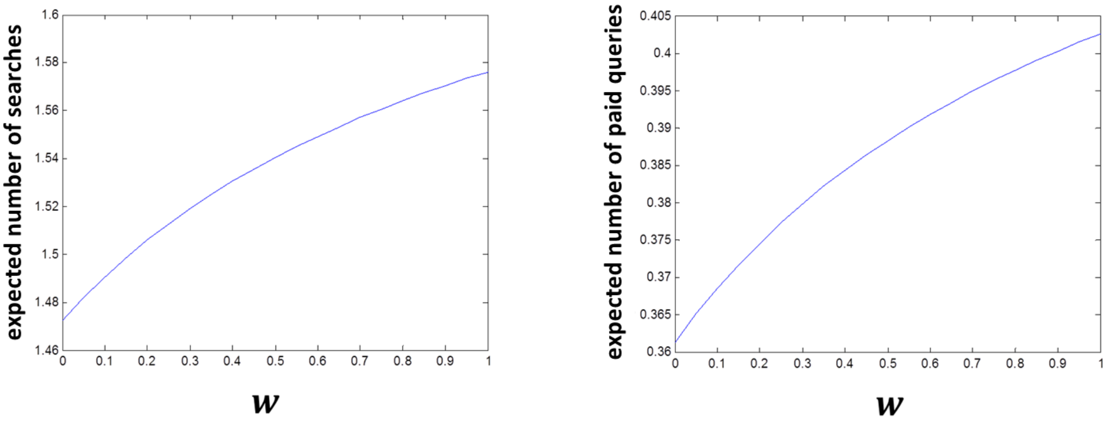


Figure 5.6: The expected number of search rounds carried out by the searcher (left) and the expected number of paid queries made by the searcher to the expert (right), as a function of w. The setting is identical to the one used for Figure 5.5.

information overall and seemingly can identify favorable opportunities more easily. Yet, at the same time the improvement in the searcher's ability to distinguish the favorable opportunities from the non-favorable ones translates to a greater expected benefit from resuming the search process, resulting in a longer search. This also explains the increase in the overall number of paid queries made to the expert. While the increase in this latter measure is beneficial for the expert, it comes with a price—the expert is also experiencing an increase in the overall number of queries it is providing for free. Therefore, supplying information for free for all signals $s > t_u$ is not beneficial and the expert should take into consideration the production cost d_e. Figure 5.7 shows the expected profit of the expert as a function of w (see Equation (5.7)). Indeed, the expected profit increases as w increases; however using $w = 1$ is not the best response strategy for the expert. The expert should offer the service for free only when the signal is such that the expected benefit from providing it (taking into consideration the chance the true value will indeed turn out to be poor and an additional search round will be initiated and the expected profit from having the searcher resume the search) is greater than the cost of providing the service for free. Formally, this is expressed as $\pi_e F_v(V|s) > d_e$ and depicted in the right graph of Figure 5.7.

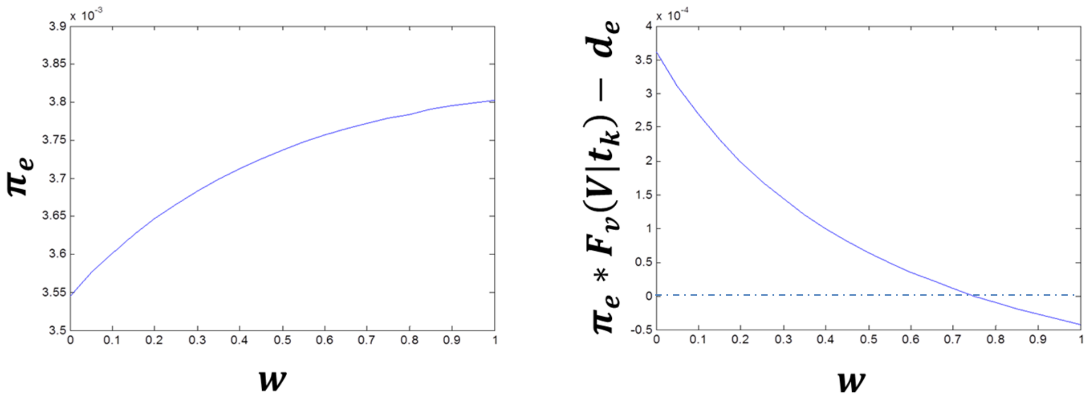


Figure 5.7: (a) The expert's expected profit from having the searcher resume search as a function of the parameter w; (b) The expected net benefit from providing the true value for free when the signal is t_k. The setting is identical to the one used for Figure 5.5. The expected net benefit in this example becomes zero for $w = 0.75$.

I emphasize that this result (both the expert and the searcher benefiting from the fact that some of the information is offered for free) is limited to the simultaneous game induced after the price is set by the expert. This does not mean that the searcher benefits overall from the expert changing its strategy to provide some information for free—at the end of the day the expert is setting c_e strategically, and it is possible that the searcher does worse overall in the world where the expert has the added flexibility to offer its services for free sometimes. For example, in the setting analyzed above (where $c_s = 0.17$ and $d_e = 0.00019$, with the expert's expected-profit-maximizing $c_e = 0.01$ (when free information disclosure is allowed) the searcher's expected benefit is 0.196, whereas when free information disclosure is not allowed the expert uses $c_e = 0.05$ and the searcher's expected benefit is 0.247.

5.5 Model Robustness

One fear for an expert or a platform when considering switching to offering a service for free is that some parts of the population that were not using the service up until then because of its price, could start using it extensively once it is offered for free, causing a substantial unexpected expense for the expert, who may not previously have been aware of their existence. In this section I illustrate numerically that even with a relatively large population of such "free riders" the expert can still benefit from offering the service for free for some signals. For this purpose I consider two populations of searchers. The first is of searchers characterized by a relatively small search cost, hence with a smaller incentive to use the expert services (as they can potentially repeat their search process until running into an opportunity associated with a very high signal, and choose to terminate the search without ever querying the expert). This population will, however, use the expert's services whenever offered for free, since this is a dominant strategy when available. The second population is characterized by a higher search cost, and uses the expert's services for some signals even when offered for a fee $c_e \gg 0$. Both populations receive signals from the same distribution $f_s(y)$ and similarly share the same function $f_v(v|s)$ according to Equation (5.4). The search costs of the two populations are $c_s^l = 0.0292$ for the low search cost population and $c_s^h = 0.17$ for the high search cost population. The expert's marginal cost for providing the service is $d_e = 1.9 \cdot 10^{-4}$ for both populations.

Based on the parameters above there is no query fee $c_e \geq d_e = 1.9 \cdot 10^{-4}$ that results in the use of the expert's services by the low search cost searchers (i.e., $t_l = t_u$ for this population). Therefore, the expert maximizes its expected-profit based on the second population only, resulting in the following equilibrium: $t_l = 0.246$, $t_u = 0.496$, $V = 0.0195$ and $c_e = 0.01$. When offering information for free for some of the signals, the expert, who cannot distinguish between searchers of the two populations, needs to take into consideration the loss due to the use of its services by searchers of the low search cost population.

Taking α to be the portion of the high search cost searchers in the general population, the expert's expected profit is given by

$$(1 - \alpha)(-d_e)\frac{F(t_k) - F(t_u)}{A} + \alpha\left((-d_e)\frac{F(t_k) - F(t_u)}{A} + (c_e - d_e)\frac{F(t_u) - F(t_l)}{A}\right),$$

where A is the probability the search is terminated (calculated according to Equation (5.5)). The first term corresponds to the loss due to the free usage of the expert's services by the low search cost searchers. The second term corresponds to the expected profit from the high search cost searchers and includes both the loss due to free service and the gain from the paid service. Both terms are weighted according to the proportion of the different searchers' types in the population.

Figure 5.8 depicts the expert's expected profit for the setting described above as a function of α, when free information disclosure is allowed and when it is not allowed. The figure demonstrates that, indeed, even for cases where the population of "free-rider" searchers is substantial (99% in this case), the expert can still benefit from free information disclosure.

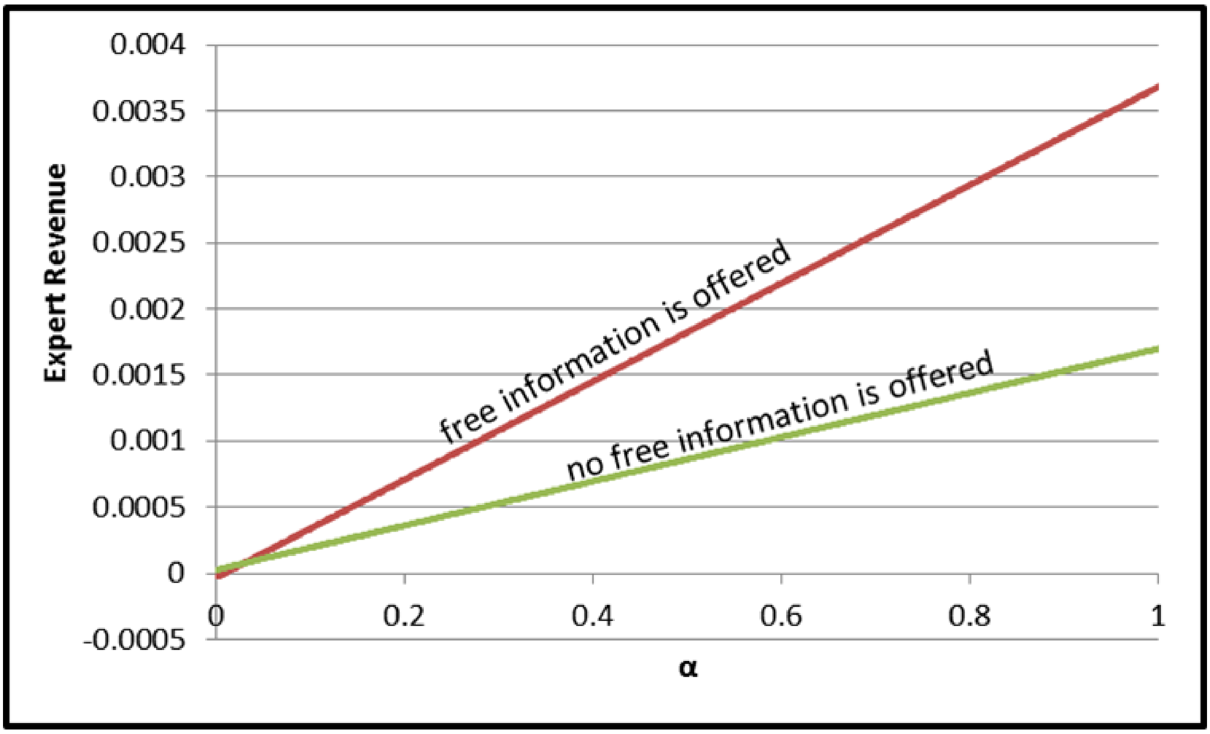


Figure 5.8: The expert's expected benefit when free information disclosure is allowed and when it is not allowed, as a function of the percentage of the high search cost searchers in the general population, for the example described in the text.

5.6 Conclusions

This chapter's main contribution is to analyze a subtle strategic complexity (free information disclosure) in a common multi-agent environment (one-sided search with a self-interested information provider or platform). The channel of operation is complex: when the expert sometimes provides its services for free, it changes the searchers' optimal strategies, expanding the range at which users choose to use its non-free services.

One natural fear in using free disclosure strategies would be model robustness – suppose the expected higher profits were driven by a misestimation of the population? For example, it could be that only those with high search costs were using expert services earlier, so the expert assumes the population in general has high search costs – however, by offering its services for free, it suddenly draws out the population with low search costs that it was unaware of previously since they never used its services. I show that the result is robust to even a significant proportion of such "free riders" in the searching population. As such, the idea of free information disclosure could have significant practical value in search-based markets and systems. I note that the information-provider in this model is working within a somewhat restricted strategy space and could have incorporated different prices (including zero, i.e. free) for each signal. Yet, one of the major results is that, even with the restricted strategy space, there is a benefit to the information provider of providing some services for free. I also note that it is important that the expert will also observe the signal s, as otherwise the searcher could lie about the signal and always get the service for free. In various domains the signal can be verified (e.g., in the used car domain, the expert (e.g., mechanic) can verify the signal by checking the Carfax report herself).

Chapter 6

Providing Information to People

This chapter [1] studies the benefit for information providers in free public information disclosure in settings where the prospective information buyers are people. The underlying model, which applies to numerous real-life situations, considers a standard decision making setting where the decision maker is uncertain about the outcomes of her decision. The information provider can fully disambiguate this uncertainty and wishes to maximize her profit from selling such information. I use a series of AMT-based experiments with people to test the benefit for the information provider from reducing some of the uncertainty associated with the decision maker's problem, for free. Free information disclosure of this kind can be proved to be ineffective when the buyer is a fully rational agent. Yet, when it comes to people I manage to demonstrate that a substantial improvement in the information provider's profit can be achieved with such an approach. The analysis of the results reveals that the primary reason for this phenomena is people's failure to consider the strategic nature of the interaction with the information provider. People's inability to properly calculate the value of information is found to be secondary in its influence.

6.1 Introduction

While the study of the use in strategic information disclosure by self interested information providers in multi-agent settings exists, the focus of prior work aiming to study strategic be-

[1]The work reported in this chapter was published in [3]

havior of such entities is mostly limited to dealing with rational agents. In this chapter I investigate the way free information disclosure influences people. The main idea of free information disclosure is that through the elimination of some of the possible outcomes, knowing the true outcome becomes highly valuable. For example, consider a passenger that is about to go on a flight from NY to Paris in order to attend an important business meeting. Now suppose the possible outcomes of the flight are: (i) arriving on time, with an a priori probability of 94.4%; (ii) arriving an hour late, with an a priori probability of 4.1%; or (iii) missing the meeting because the flight gets canceled due to a union strike, with an a priori probability of 1.5%. Knowing the true outcome (e.g., by purchasing it from an oracle or a corrupted union member) has very little value, as the chance of not arriving to the meeting on time is very small (1.5%). However, assume the oracle publicly announces that the flight is not going to arrive on time (i.e., eliminating the first outcome, hence reducing the set of possible outcomes to the latter two). Now, there is much value in being able to distinguish between the two remaining outcomes—the naive posterior probability of ending up with a canceled flight due to a strike is 27%. Therefore, the passenger will be willing to pay a substantial amount in order to obtain this information. Still, the above naive probability update process does not take into account the strategic considerations that lead the information provider to publicly disclose some of the information she holds. The incorporation of the strategic aspect of the interaction results in a somehow different probabilistic update and in fact I can prove that free information disclosure is necessarily detrimental in this case. However, when dealing with people, the above does not necessarily hold. It is well known that people are often irrational [94, 65, 10, 23]. Therefore, it is possible that they will not take into consideration the strategic nature of the interaction or even fail to properly reason about the value of information to some extent, making free information disclosure beneficial for the information provider.

This chapter provides a comprehensive experimental evaluation of the above approach whenever interacting with people, attempting to identify the main sources of people's inability to make the right decision when it comes to information purchasing. It uses a testbed that captures a core “value of information" problem setting of the kind described above, with an information provider that can fully disambiguate the uncertainty associated with outcomes. The experiments involve 300 subjects experiencing a total of 6000 information purchasing decisions, interacted through Amazon Mechanical Turk using three different treatments.

Contributions. This chapter makes two main contributions. The first is showing that, unlike with fully rational buyers, when it comes to people free information disclosure can substantially improve the information provider's profit from selling information. The second is showing that the improvement achieved is mostly because of people's inability to take into consideration the strategic nature of the interaction rather than their somehow limited ability to properly calculate the value of information.

6.2 The Model

I consider the basic standard model of a self-interested information provider and a prospective information buyer (denoted “buyer" onwards). The buyer is facing a simple decision problem involving an opportunity O available to her, where the possible available alternatives are to exploit opportunity O or opt-out not to exploit it. The set of possible exploitation outcomes (corresponding to different possible nature states) is denoted $V = \{v_1, v_2, ..., v_n\}$, where the corresponding a priori probability of each value $v \in V$ is captured by the function $p(v)$ ($\sum p(v_i) = 1$). If choosing to opt-out, the buyer gains some fallback profit $v_\emptyset$. The buyer and the information provider are symmetric in the sense that they are both familiar with V and the function $p(v)$. The information provider is also acquainted with the true state of the world, i.e., knows the true exploitation value of O and can sell this information to the buyer for a fee. In an effort to increase her profit, the information provider can use a strategic behavior and publicly eliminate some of the possible outcomes of O such that this information becomes available to the buyer before she makes her decision of whether to purchase the identity of the true outcome or not. I denote this latter strategy PFID (Preliminary partial Free Information Disclosure) for short.

The course of the game is therefore as follows: nature first sets the true exploitation outcome v of the opportunity O; the value v becomes available to the information provider who sets the requested fee c for revealing v along with eliminating (publicly) some of the values in V, such that the remaining possible values are those in the subset $D \subseteq V$; based on c and D, the buyer can decide either to purchase the true exploitation value of O, in which case the value v is disclosed to her, or not; finally, the buyer decides whether to exploit opportunity O.

The model assumes that both the exploitation values and the cost of purchasing the information from the information provider are additive. The goal of the buyer is therefore to maximize her expected profit, defined as the exploitation value of O (if exploiting the opportunity) or the fallback $v_\emptyset$ (otherwise) minus the payment to the information provider (if purchasing the information).

The above model can be mapped to various real-life problems. For example, the buyer can represent a company that considers taking over its competitor. The true value of the other company is uncertain however can be purchased from an internal source that may increase the value of the information she holds through PFID. The information provider's problem is thus, given an opportunity O and the true exploitation value v, which exploitation values to eliminate for free and what price to set as the fee for revealing v in order to maximize her expected profit.

6.3 Rational Buyers

I first analyze the best response strategies and the resulting equilibrium in case the buyer is fully rational and risk neutral.

Buyer

In the absence of any preliminary information from the information provider, the buyer will choose to exploit opportunity O only if the expected exploitation value is greater than the fallback value $v_\emptyset$. The buyer's **Expected Monetary Value (EMV)** is thus given by:

$$EMV(O) = max(\sum_{v \in V} v \cdot p(v), v_\emptyset) \tag{6.1}$$

If purchasing the information from the information provider, the buyer's decision is made under certainty. Here, the buyer exploits O only in cases where the exploitation value is greater than the fallback utility. The **Expected Value Under Certainty (EVUC)** is thus given by:

$$EVUC(O) = \sum_{v \in V} max(v, v_{\emptyset}) \cdot p(v) \tag{6.2}$$

The **Value of the Information** held by the information provider to the buyer (denoted VoI onwards) is thus $VoI(O) = EVUC(O) - EMV(O)$ and this is the maximum amount the buyer will be willing to pay for receiving the true outcome.

Finally, when the information provider uses PFID, leaving only a subset D of remaining applicable outcomes, the above calculations still hold with some minor modifications:

$$\begin{aligned} &VoI(D) = EVUC(D) - EMV(D) = \\ &\sum_{v \in D} max(v, v_{\emptyset}) \cdot Pr(v|D) - max(\sum_{v \in D} v \cdot Pr(v|D), v_{\emptyset}) \end{aligned} \tag{6.3}$$

where $Pr(v|D)$ is the posterior probability of the exploitation value being v given the evidence D. Naively, the value of $Pr(v|D)$ should be calculated through a simple update of the a priori probability $p(v)$ as follows:

$$Pr(v|D) = \begin{cases} \frac{p(v)}{\sum_{y \in D} p(y)} & \text{if } v \in D \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

The above calculation is considered naive as it does not take into consideration the strategic behavior of the information provider. Recall that the information provider's strategy is a function $S : V \to D$ (and the corresponded prices to be charged for revealing the true value, calculated as the VoI), specifying for each outcome $v \in V$ the subset $D \subseteq V$ of remaining possible exploitation values. The strategy thus induces a partition of the set V such that any two outcomes v_i and v_j are in the same partition element if and only if $S(v_i) = S(v_j)$. Since in equilibrium the buyer is using her best response strategy to the information provider's strategy, the posterior probabilities calculation taking place by the buyer should be based on S and is given by:

$$Pr(v|D) = \begin{cases} \frac{p(v)}{\sum_{y|S(y)=D} p(y)} & \text{if } S(v) = D \\ 0 & \text{otherwise} \end{cases} \tag{6.5}$$

Information Provider

As explained above, the information provider sets the cost of her information providing service to VoI. The information provider may attempt to maximize the VoI through PFID. In this case I distinguish between having a naive buyer and a strategic one.

Naive Buyer When the buyer does not take into consideration the fact that the information provider is acting strategically she uses the naive probability update according to (6.4). For example, assume that $V = \{-100, 0, 100\}$ where all values are possible with equal probability, and assume the fallback if not exploiting the opportunity is $v_\emptyset = 0$. Here VoI=33.3. However, if the information provider uses $S(-100) = S(100) = \{-100, 100\}$ and $S(0) = \{-100, 0, 100\}$, she can still charge 33.3 whenever $v = 0$ however charge 50 in case $v \in \{-100, 100\}$. Generally, when facing a naive buyer, the information provider should choose for every value $v \in V$ to eliminate all exploitation values except those in $D \subseteq V$ such that the difference $EVUC(D) - EMV(D)$ is maximized and charge exactly the difference.

Strategic Buyer When the buyer is acting strategically, however, the information provider cannot benefit from free information disclosure, as stated in Proposition 3.

Proposition 3. *The information provider's expected profit when using PFID is bounded (from above) by the expected profit when not using it.*

Proof. Since the information provider sets the price of her service according the worth of the information, I need to show that the expected VoI under certainty without PFID is greater than with PFID. Meaning that the following holds:

$$EVUC(O) - EMV(O) \geq \sum_{D} (EVUC(D) - EMV(D)) \cdot Pr(D) \tag{6.6}$$

where $Pr(D)$ is the probability the buyer will receive the information D, calculated according to: $Pr(D) = \sum_{v \in D} p(v)$. Notice that:

$$\begin{aligned} EVUC(O) &= \sum_{v \in V} max(v, v_\emptyset) \cdot p(v) = \sum_{D} Pr(D) \sum_{v \in D} max(v, v_\emptyset) \cdot Pr(v|D) \\ &= \sum_{D} EVUC(D) \cdot Pr(D) \end{aligned}$$

therefore, in order for (6.6) to hold I only need to prove that $\sum_D EMV(D) \cdot Pr(D) \geq EMV(O)$:

$$
\begin{aligned}
& EMV(O) = max\big(\sum_{v \in V} v \cdot p(v), v_{\emptyset}\big) = \\
& max\big(\sum_{D} Pr(D) \cdot \sum_{v \in D} v \cdot Pr(v|D), v_{\emptyset}\big) \leq \\
& \sum_{D} Pr(D) \cdot max\big(\sum_{v \in D} v \cdot Pr(v|D), v_{\emptyset}\big) = \\
& \sum_{D} EMV(D) \cdot Pr(D) \quad \square
\end{aligned}
\tag{6.7}
$$

Therefore, if both the information provider and the buyer are fully rational and strategic, there is no point for the information provider to use PFID. In the following section, however, I show experimentally that there is much value in such strategy when the buyer is a person.

6.4 Irrational Buyers

In most real-world settings one expects to find people in the role of the buyer. This section describes an experiment carried out for testing the effectiveness of PFID in such a case.

6.4.1 Possible Failures in Decision Making

Prior work provides much evidence for people's bounded rationality in decision making situations in the sense that they do not adhere to rigid models of rationality and are easily influenced by various external factors and biased towards certain conclusions [109, 78, 65, 53, 73]. Specifically, for the strategic interaction settings considered in this chapter I identify two possible causes for irrational behavior that may affect the decision whether to purchase the information offered. The first is people's somehow limited reasoning and computational capabilities that may prevent the proper calculation of the value encapsulated in the information according to the guidelines given in the former section. The second is people's failure to take into consideration the strategic nature of the interaction with the information provider. The implication of the latter is failure to update the probabilities assigned to the different exploitation values according to (6.5) and using the following naive calculation instead:

$$Pr(v|D) = \begin{cases} \frac{p(v)}{\sum_{y \in D} p(y)} & \text{if } S(v) = D \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

Meaning that the buyer does not take into consideration the reason the information provider decided to disclose D rather than any other subset.

I note that prior literature contains evidence for both above phenomena, i.e., people's failure to take into consideration the strategic aspect of an interaction [35] and failure to accurately calculate the value of information [66, 18]. The extent of the effect, if any, depends on the domain, the nature of the interaction and the complexity of the underlying problem. Still, none of these works consider a model similar to this one and the results reported there cannot be trivially carried over to this case.

Among the two effects, the second clearly favors the use of PFID in a way that increases the value of the information held by the information provider whenever the buyer follows (6.4). The effect of the inability to properly calculate the value of information (i.e., even if taking into consideration the strategic aspect of the interaction) when using PFID is somehow vague, as it is not clear whether it will actually result in an increase or a decrease in the value buyers see based on the information provided, even in cases where VoI(D) increases. The experiments were designed such that both effects can be isolated to a great extent.

Experimental Framework

For the experiments I used a multi-round game called **"What's In The Box?"**, which captures the essence of the basic underlying decision making problem in the model without adding any externalities that may confuse participants. On each round in the game the player is introduced with a box which contains a prize expressed in game points (corresponding to an opportunity in the model). The available alternatives are to open the box (corresponding to exploiting it) or leave it unopened (opt-out). Along with the box the player is also introduced with the possible values of the prize in it (corresponding to the possible exploitation values). Prize values can be either positive or negative, each having an a priori equal chance. Prior to her decision whether to open the box, the player can request to obtain the identity of the prize in the box, i.e., completely disambiguate the uncertainty associated with the value. This latter information is, however, costly, and the cost of obtaining it (expressed in terms of game points) is provided to the player prior to making her decision to request it. The player thus needs to decide whether to purchase the information about the true value of the prize in the box and then whether to open the box. If choosing not to open the box the player obtains zero game points (the fallback value). Finally, the player moves on to the next game round, and the appropriate adjustments to her total accumulated game points are made (adding the prize (or actually reducing it in case its value is negative) in case the box was opened and reducing the cost of information if purchased). The goal of the player is to accumulate as many game points as possible throughout the game.

I note that the primary reason for choosing a repeated game where on each round the player is facing a different decision problem instance (though of similar nature) was to have people follow an EMV-based decision rule. Prior work provides much evidence for the fact that in repeated-play settings people's strategies asymptotically approach the EMV strategy as the number of repeated plays increases [72, 69, 17]. The proper solution to the game, when taking an EMV-maximizing approach is quite straightforward and follows exactly the calculation given in the section dealing with rational buyers: the player should purchase the information if $EVUC(O) - EMV(O)$ (or $EVUC(D) - EMV(D)$ when using PFID) is greater than its cost, and open the box only if the value of the prize (or the expected value of the prize in case the information is not purchased) is greater than zero.

Experimental Design

I implemented the "What's In The Box?" game using $C\#.net$ for the server side and $Html5$, css and $Javascript$ for the client side such that participants could interact with the system using a relatively simple graphic interface. Figure 6.1 present a screen shot of the game where the player is introduced to the possible values of the prize in the box (all with the same probability) and the cost of purchasing the true value. In this stage, the buttons provided to the player to make her decision are disabled for the first ten seconds so she is forced to spend some time thinking before making her decision.

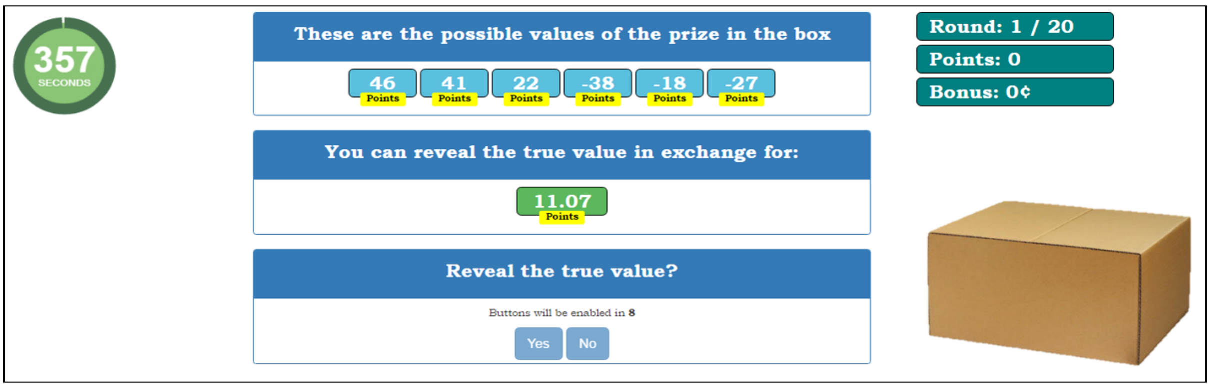


Figure 6.1: Screen shot of the game. See text for details.

In order to support free information disclosure, I enabled crossing out some of the possible values of the prize in the box few seconds after they appear, so that the player could still see the set of original values and those that have been removed. At the end of each round the player received a short summary detailing the change in her accumulated game points, listing the prize obtained (if opening the box) and the payment for the information (if purchased).

I used three different experimental treatments:

No Free Information Disclosure - where no free information disclosure takes place, i.e., none of the values is crossed out prior to the information purchase decision.

Free Information Disclosure by an Explicitly Strategic Information Provider - where information is sold by a strategic information provider that uses PFID. With this treatment I did everything possible, from the UI point of view, to make sure the player understands that values are being eliminated by a self-interested agent that aims to maximize her own gain. Therefore the player was told that there is additional player in the game, who gains from selling the information to her. In each round, in addition to presenting the player's own accumulated score on the screen, I also presented the information provider's accumulated profit.

Free Information Disclosure by a Non Explicitly Strategic Information Provider - where information is sold by a strategic information provider, except that with no mentioning of the strategic considerations accounting for the disclosure of information. Participants were told that values are removed by the "system" as a way of helping the player and obviously there was no mentioning or reflection of the information provider (or her score) in the GUI. The idea in including this treatment in the experiment was to see how close will be the decisions of players under this treatment to those exhibited in the second treatment. A great similarity would indicate that people tend to ignore the strategic nature of the information provider in the second treatment. The use of PFID in the last two treatments followed the guidelines provided at the end of the Naive buyer part in the Rational-buyers section above (i.e., maximizing the expected profit assuming facing a naive buyer).

In order to have better control over the experiment I pre-generated a set of core problem settings. The values for the different outcomes in each problem were integers randomly picked within the range $[-50, 50]$. In order to reason about the effect of the number of values on the results obtained I generated a total of 250 such problems, differing in their number of outcomes n, in a way that I had 50 problems for each number of outcomes $n \in [3, 7]$. In order to reason about the effect of the magnitude of the difference between the value of information and its cost on people's ability to make the right decision, I took the cost to be exogenously set (rather than setting it as VoI(O)).[2] For this purpose I created four problem instances based on each core setting O (of the 250 mentioned) differing in the cost of purchasing the information, setting the price of information to: (1) $0.8\cdot$VoI(O); (2) $1.2\cdot$VoI(O); (3) $0.2\cdot$VoI(O); and (4) $1.8\cdot$VoI(O). In those few cases where VoI(O)=0 (e.g., when all outcomes are positive) I randomly picked the cost of information for each of the four resulting problem instances (within the range [0,50]). The full set of problem instances is available upon request. Overall, in 48% of the problems a rational buyer should purchase the information and in the remaining 52% she should not, where the difference corresponds to those cases where VoI(O)=0 (hence information should not be purchased regardless of its price).

Participants were recruited and interacted through Amazon Mechanical Turk (AMT). Although AMT consider to be biases in some cases (for example since the service, so far, is only available in English and to make job requests you have to have a U.S. address), it is already been proven to be a well established method for data collection in tasks which require human

[2] As otherwise, if the information provider sets the price to be exactly VoI(O) even the slightest deviation in the calculation of the value of information may lead to wrong results, precluding a genuine analysis of the extent to which people are affected by their failure to take the information provider as a strategic agent.

intelligence [90]. To prevent any carryover effect a “between subjects" design was used, assigning each participant to one treatment only. The compensation for taking part in the experiment was composed of a show-up fee (the basic “HIT") and also included a bonus, which was a direct outcome of the participant’s performances in the experiment (measured as the amount of accumulated game points), in order to encourage thoughtful participation—one cent bonus for each 10 game points accumulated. Each participant received thorough instructions of the game rules, the compensation terms and her goal in the game. Then, participants were asked to engage in practice games until stating that they understood the game rules (with a strict requirement for playing at least two practice games). Prior to moving on to the actual games, participants had to correctly answer a short quiz, making sure they fully understand the game and the compensation method. Finally, participants were requested to play a sequence of 20 rounds, where the problem instance used for each round was randomly picked from the pool of 1000 problem instances described above (with no repetition).

During the game, I logged all player actions along the different phases (instructions, training, quiz and actual game). I had four classifications for each player’s information purchasing decision: whenever purchasing the information, the decision was classified as “good" if the VoI is greater than or equal to its cost (and “bad" otherwise). Similarly, whenever not purchasing, the decision was classified as “good" if the VoI is lower than or equal to its cost (and “bad" otherwise). The above was calculated in all three treatments according to the naive VoI calculation as described in (6.8). For the two treatments that use PFID, I repeated the calculation by taking the VoI to be calculated according to (6.5) assuming the information provider applies the PFID as described above.

Results

Overall, I had 300 participants taking part in the experiments, 100 for each experiment, each playing 20 rounds according to the above design. Participants ranged in age (18-81, average 34.5) and gender (57% men and 43% women), with a fairly balanced division between treatments.

The analysis of the results shows that the virtual information provider managed to substantially improve the overall profit from selling information when using PFID, compared to when not using it. The following table details the average per-game (20 rounds) profit obtained with each of the three treatments and the number of instances in which the player purchased the information.

	Treatment 1	Treatment 2	Treatment 3
Avg.Total Profit	57.2	73.6	77.4
# of sales	1030	1206	1189

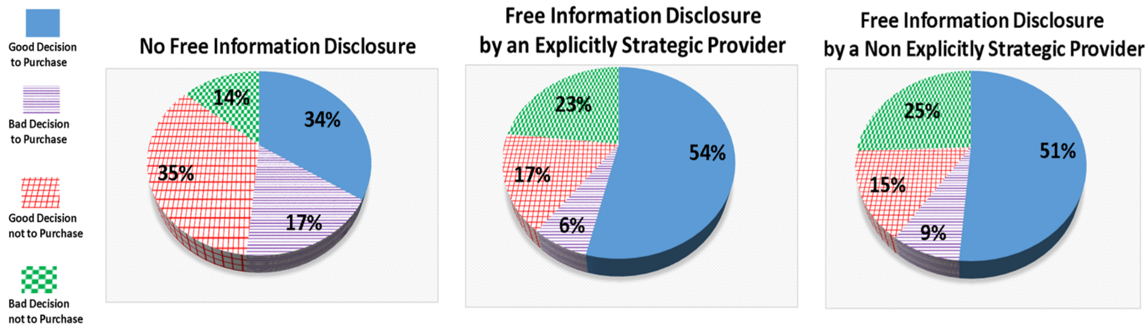


Figure 6.2: Classification (using naive VoI calculation) of decisions made in all treatments.

The above table reflects an increase of 29% and 35% (both statistically significant using $t-test$, $p < 0.005$) in the information provider's profit through PFID (compared to when not using it), when presenting the information provider as a fully strategic player (treatment 2) and when avoiding any mentioning of her strategic nature (treatment 3), respectively. The improvement in the information provider's expected profit due to not presenting her as a strategic player (i.e., in the transition from treatments 2 to 3) is 5% (non statistically significant using $t-test$, $p > 0.5$). Similarly, the number of instances in which the information provider managed to sell the information she was holding when using PFID (i.e., (treatments 2 and 3), increased by 17% and 15% (both statistically significant using $t-test$, $p < 0.005$) with the two information-disclosure treatments (and a minor reduction of 1% in the transition in-between the last two (non statistically significant using $t-test$, $p > 0.5$)). The insignificant differences between the profits obtained with treatments 2 and 3, as well as further similarities observed in the in-depth analysis of the results, as reported in the following paragraphs, suggest that people do not take into consideration the strategic aspect of the problem they are facing in this domain.

One additional evidence that strengthens this latter hypothesis can be found in the performance achieved in the third treatment. If the players were fully rational as far as the computation of the VoI is concerned, yet still naive in the sense of not taking the information provider to be strategic, then the theoretical expected profit of the information provider based on the 1000 problem instances is 85.76. The information provider in the third treatment, the one that emulates this exact scenario, managed to reach a very close profit (77.4).

Figure 6.2 provides a more detailed investigation concerning the sources of the improvement achieved with PFID. It depicts the break-down of the total 2000 information purchase decisions made in each treatment into the four different classifications described in the experimental design section (based on naive VoI calculation). Considering the chart that summarizes the results obtained when not using PFID (most left), I observe that the general success of people with the tested settings is 69%, with relatively similar chance of choosing the wrong action according to the two classifications (either purchase when better not to purchase and vice versa). These latter findings suggest that people are indeed unable to properly calculate the value of information to some extent. With the information providers using PFID I observe that the percentage of instances in which information is purchased increases from 51% to 60% (regardless of how the information provider was presented to the players). Interestingly, in the second and third treatments the percentage of cases where information was purchased out of those where it should not had been ($23/54 = 30\%$ and $25/76 = 33\%$, respectively) or when not purchased out of those where it should had been (26%,37%) did not change much between treatments. This, as well as the relatively similar division into the four classifications observed within the charts corresponding to the second and third treatments, indicate, once again, that it is people's failure to consider the information provider to be strategic that accounts for most of the improvement achieved in the information provider's profit. The inability to accurately calculate the value of information is definitely reflected in the results, as explained above, yet its impact is only secondary to the primary effect.

Figure 6.3 depicts the average success rate of players in their decision of whether to purchase the information from the information provider, as a function of the benefit in purchasing it (VoI-cost) in the different treatments. The success rate is measured as the percentage of de-

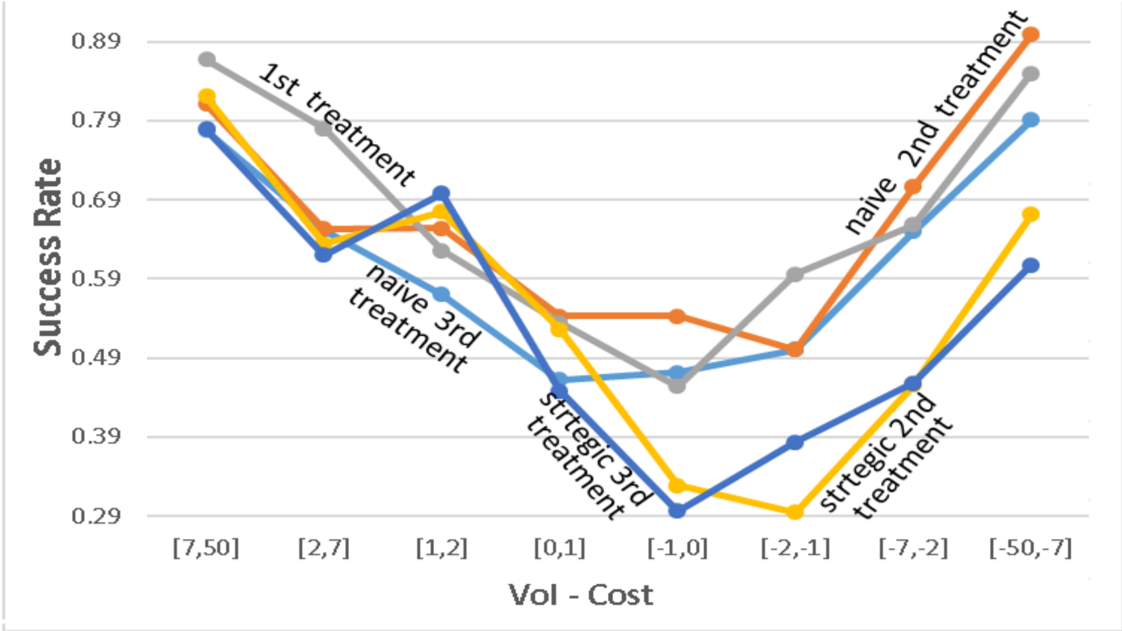


Figure 6.3: Players' success rate in the different treatments.

cisions classified as "good" out of all those made. For each of the two treatments using PFID I included two curves. The first refers to classifications according to the naive calculation of the VoI and the second according to the calculation that takes all strategic considerations into account. From the graph one can observe that people are quite good in realizing that purchasing the information is beneficial (or not beneficial) whenever the difference between the true value of the information and its cost is substantial. The greater the difference, the better the quality of the decision people make.

The fact that the two curves corresponding to the naive information value calculation under PFID almost entirely coincide with the curve corresponding to not using PFID suggests that people completely fail to consider the strategic behavior of the information provider. The improvement achieved in the information provider's profit is thus primarily through the increase in the number of instances where the value of information becomes greater than its cost. Indeed, even with PFID people still reflect the same computational difficulties in reasoning about the benefit in purchasing the information, however since the overall number of "beneficial" instances increases so does the number of times information is purchased. The two curves representing the quality of people's decisions when the VoI calculation takes strategic considerations into account are very close. Their general behavior also reflects better success whenever the benefit in purchasing the information is relatively high or low, though their center point is shifted compared to the others. These two insights complement all the findings reported so far related to the role of the two hypothesized reasons in generating the benefit PFID achieves.

6.5 Conclusions

The encouraging results reported in the results section suggest that information providers can greatly benefit from free information disclosure when facing human buyers. The importance of this finding is primarily due to the fact that real-world information buyers are human (as opposed to fully rational agents), the extensive penetration of strategic information providers to almost any field in our lives and the wide applicability of the underlying decision making model used. These results will be valuable both for practitioners developing information providing platforms and applications and for researchers who hopefully will see the potential in continuing this line of work and design and test more advanced methods for improving information providers' revenues when interacting with people, based on the insights provided in this chapter.

The results' analysis unfolds the main reason for the success of the proposed approach: it is primarily people's failure to consider the strategic nature of the interaction that precludes a proper judgment. Therefore, an increase in the naive value of certainty translates to an almost identical increase in its value in the eyes of the buyer.

Chapter 7

Final Remarks

In this thesis, the role of information's effective providing has been investigated in a number of multi-agent systems. The different characteristics of each system presented here allow the examination of various uses of information as part of exploring the most efficient way to provide information. In a way, this is the first step toward a theory of effective information providing for platforms, information providers and market designers.

Many of the results given in the thesis are somehow counter-intuitive and offer new paradigms for information providing. For example, in auctions I was able to show that by augmenting the information provider's expected profit and allowing her to disclose part of the information she holds for free (in addition to setting the price for which she is willing to sell the information), the information provider's expected profit can be increased. Moreover, I was also able to show that due to the use of free information disclosure, the whole system's expected profit (i.e., the social welfare) can be increased. I was able to show that in some cases the auctioneer's expected profit will increase when receiving less accurate information from the information provider. As a result, she will sometimes even take some actions in order to prevent the information provider from fully identifying the exact state of the world. Finally, even when preventing the auctioneer and the bidders from knowing the identity of the player who discloses the information, the information provider can still benefit from the information disclosure, taking advantage of the fact that the bidders and the auctioneer do not take her to be a strategic player.

For one-sided economic search settings, I was able to show that when agreeing to provide some of the information for free (i.e., for a specific range of signals) the information provider/platform can actually increase her profit from the search, pushing the searcher to use her costly services for a wider range of signals. I was also able to show that disclosing part of the information for free may also benefit the expected profit of the searcher from the search (e.g., when dealing with a "lemon" opportunity). Second, I was able to characterize the structure of the equilibrium for a model where the information provider discloses part of the information for free, which can be very useful for an information provider/platform when deciding on her information providing strategy. Finally, I was able to show that the results achieved are robust to even a significant proportion of "free rider" agents (i.e., agents who only use the free services provide by the information provider/platform) in the searching population. Therefore, the idea of free information disclosure could have significant practical value in search-based markets and systems.

There are several interesting extensions of the proposed work presented in Chapters 3 and 4. One natural extension is relaxing the assumption that the decision regarding purchasing the information is exclusively the auctioneer's, allowing bidders to purchase the information directly from the information provider. Indeed, the original assumption holds in some real-world situations, e.g., when the information provider's services might require direct access to the auctioned item or some information that the auctioneer holds. However, in many others there is a good reason to believe that the information can be purchased also by the bidders. Note that such models, where bidders also have direct access to the information, need to be carefully dealt with and analyzed, since there are many important modeling choices that need to be made (for example: Can the information be sold to more than a single bidder? Will the auctioneer be able to purchase the information? Will those purchasing the information be able to disclose it to any of the other players? Will the other players know who purchased the information? Will the information be offered for sale sequentially or to all players in parallel? Can the information provider set a different price for different players (e.g., to the auctioneer and to the bidders)?). All of these choices will certainly affect the analysis and the nature of the dynamics formed. An additional natural extension is the study of multi-information-provider competition which can benefit greatly from the analysis provided in these chapters.

Similarly, various fascinating extensions to the work presented in Chapter 5 are available. One interesting extension is studying the non-verifiable signals domain (in contrast to the model studied in Chapter 5 for which it is important that the expert verify the signal given by the searcher). Other important avenues for future work include analysis of information provision and the incentives for free reporting in two-sided search markets (for example, matchmakers in a dating service) and analysis of searchers' incentives for truthfulness. In Chapter 5, I assume that searchers truthfully disclose their signals, which makes sense in verifiable settings like presenting a Carfax report to a mechanic; however, in more subjective settings like dating or travel preferences, how can the expert guarantee that the user is revealing her signal truthfully?

Finally, by performing a series of AMT experiments I was able to show that in contrast to the theoretical results, when providing information to people, the information provider can benefit from using partial free information disclosure. Furthermore, I showed that the main reason for the improvement in the information provider's expected profit is the fact that people do not consider her to be a strategic player. The fact that people are struggling with the calculation of the correct value of the information has a smaller influence on the information provider's expected profit. An important extension of this work is the one where the information provider also has the price-setting capability (i.e., is able to set the price of the information in addition to deciding which information to disclose for free). Having control over both the price asked and the information disclosed prior to the buyer's decision whether to purchase the information can certainly increase the information provider's profit. Still, this requires learning the mutual effects of these two parameters, as irrational buyers may be affected by different combinations of price and disclosed information in various ways.

Bibliography

[1] G. Akerlof. The market for “lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

[2] S. Alkoby and D. Sarne. Strategic free information disclosure for a vickrey auction. In *International Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis*, pages 1–18. Springer, 2015.

[3] S. Alkoby and D. Sarne. The benefit in free information disclosure when selling information to people. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pages 985–992, 2017.

[4] S. Alkoby, D. Sarne, and S. Das. Strategic free information disclosure for search-based information platforms. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems(AAMAS-15)*, pages 635–643, 2015.

[5] S. Alkoby, D. Sarne, and E. David. Manipulating information providers access to information in auctions. In *Technologies and Applications of Artificial Intelligence*, pages 14–25. Springer, 2014.

[6] S. Alkoby, D. Sarne, and I. Milchtaich. Strategic signaling and free information disclosure in auctions. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pages 319–327, 2017.

[7] M. Armstrong. Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691, 2006.

[8] I. Ashlagi, M. Braverman, and A. Hassidim. Stability in large matching markets with complementarities. *Operations Research*, 62(4):713–732, 2014.

[9] A. Azaria, Y. Gal, S. Kraus, and C. V. Goldman. Strategic advice provision in repeated human-agent interactions. *Autonomous Agents and Multi-Agent Systems*, 30(1):4–29, 2016.

[10] A. Azaria, Z. Rabinovich, C. V. Goldman, and S. Kraus. Strategic information disclosure to people with multiple alternatives. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):64, 2015.

[11] A. Azaria, Z. Rabinovich, S. Kraus, C. V. Goldman, and O. Tsimhoni. Giving advice to people in path selection problems. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-12)*, pages 459–466, 2012.

[12] R. Azoulay and E. David. Truthful and efficient mechanisms for website dependent advertising auctions. *Multiagent and Grid Systems*, 10(2):67–94, 2014.

[13] R. Azoulay-Schwartz and S. Kraus. Stable repeated strategies for information exchange between two autonomous agents. *Artificial Intelligence*, 154(1):43–93, 2004.

[14] R. Azoulay-Schwartz, S. Kraus, and J. Wilkenfeld. Exploitation vs. exploration: choosing a supplier in an environment of incomplete information. *Decision support systems*, 38(1):1–18, 2004.

[15] A. Bagnall and I. Toft. Autonomous adaptive agents for single seller sealed bid auctions. *In Proceedings of the fifth International Conference on Autonomous Agents and Multiagent Systems (AAMAS-06)*, 12(3):259–292, 2006.

[16] J. Y. Bakos. Reducing buyer search costs: Implications for electronic marketplaces. *Management science*, 43(12):1676–1692, 1997.

[17] B. Barron. When smart groups fail. *The journal of the learning sciences*, 12(3):307–359, 2003.

[18] M. H. Bazerman and D. A. Moore. Judgment in managerial decision making. *Judgment in Managerial Decision Making*, 2008.

[19] S. Board. Revealing information in auctions: the allocation effect. *Economic Theory*, 38(1):125–135, 2009.

[20] J. Bredin, D. Parkes, and Q. Duong. Chain: A dynamic double auction framework for matching patient agents. *Journal of Artificial Intelligence Research (JAIR)*, 30:133–179, 2007.

[21] P. Bro Miltersen and O. Sheffet. Send mixed signals: earn more, work less. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 234–247. ACM, 2012.

[22] S. Bronfman, N. Alon, A. Hassidim, and A. Romm. Redesigning the israeli medical internship match. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 753–754. ACM, 2015.

[23] C. Buntain, A. Azaria, and S. Kraus. Leveraging fee-based, imperfect advisors in human-agent games of trust. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI-14)*, pages 916–922, 2014.

[24] B. Caillaud and B. Jullien. Chicken & egg: Competition among intermediation service providers. *RAND Journal of Economics*, 34:309–328, 2003.

[25] H. K. Cheng and G. J. Koehler. Optimal pricing policies of web-enabled application services. *Decision Support Systems*, 35(3):259–272, 2003.

[26] M. Chhabra, S. Das, and D. Sarne. Competitive information provision in sequential search markets. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-14)*, pages 565–572, 2014.

[27] M. Chhabra, S. Das, and D. Sarne. Expert-mediated sequential search. *European Journal of Operational Research*, 234(3):861–873, 2014.

[28] S. Das and E. Kamenica. Two-sided bandits and the dating market. In *IJCAI*, volume 5, page 19, 2005.

[29] S. Das and J. N. Tsitsiklis. When is it important to know you've been rejected? A search problem with probabilistic appearance of offers. *Journal of Economic Behavior and Organization*, 74:104–122, 2010.

[30] E. David, R. Azoulay-Schwartz, and S. Kraus. Bidding in sealed-bid and english multi-attribute auctions. *Decision Support Systems*, 42(2):527–556, 2006.

[31] E. David and E. Manisterski. Strategy proof mechanism for complex task allocations in prior consent for subtasks completion environment. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02*, pages 209–215. IEEE Computer Society, 2013.

[32] S. Dobzinski and N. Nisan. F-unit auctions. *Journal of Artificial Intelligence Research (JAIR)*, 37:85–98, 2010.

[33] M. Dufwenberg and U. Gneezy. Information disclosure in auctions: an experiment. *Journal of Economic Behavior & Organization*, 48(4):431–444, 2002.

[34] S. Dughmi, N. Immorlica, and A. Roth. Constrained signaling for welfare and revenue maximization. *ACM SIGecom Exchanges*, 12(1):53–56, 2013.

[35] K. M. Eisenhardt and M. J. Zbaracki. Strategic decision making. *Strategic management journal*, 13(S2):17–37, 1992.

[36] A. Elmalech, D. Sarne, and B. J. Grosz. Problem restructuring for better decision making in recurring decision situations. *In Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*, 29(1):1–39, 2015.

[37] Y. Emek, M. Feldman, I. Gamzu, R. PaesLeme, and M. Tennenholtz. Signaling schemes for revenue maximization. *ACM Transactions on Economics and Computation*, 2(2):5, 2014.

[38] P. Eső and B. Szentes. Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731, 2007.

[39] D. S. Evans and R. Schmalensee. The industrial organization of markets with two-sided platforms. Technical report, National Bureau of Economic Research, 2005.

[40] T. S. Ferguson. Who solved the secretary problem? *Statistical Science*, 4(3):282–289, 1989.

[41] Forbes. The world's biggest public companies. `https://www.forbes.com/global2000/list/`, 2017.

[42] W. E. Forum. The 12 biggest technology acquisitions of all time. `https://www.weforum.org/agenda/2015/10/the-12-biggest-technology-acquisitions-of-all-time/`, 2016.

[43] J.-J. Ganuza. Ignorance promotes competition: an auction model with endogenous private valuations. *RAND Journal of Economics*, pages 583–598, 2004.

[44] J.-J. Ganuza and J. S. Penalva. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030, 2010.

[45] B. Gerkey and M. Mataric. Sold!: auction methods for multirobot coordination. *IEEE Transactions on Robotics*, 18(5):758–768, 2002.

[46] J. Gilbert and F. Mosteller. Recognizing the maximum of a sequence. *Journal of the American Statistical Association*, 61:35–73, 1966.

[47] J. K. Goeree and T. Offerman. Competitive bidding in auctions with private and common values. *The Economic Journal*, 113(489):598–613, 2003.

[48] M. F. Gorman, W. D. Salisbury, and I. Brannon. Who wins when price information is more ubiquitous? An experiment to assess how infomediaries influence price. *Electronic Markets*, 19(2-3):151–162, 2009.

[49] A. Hagiu. Two-sided platforms: Pricing and social efficiency. *Mimeo, Princeton University*, 2004.

[50] A. Hagiu and R. S. Lee. Exclusivity and control. *Journal of Economics & Management Strategy*, 20(3):679–708, 2011.

[51] A. Hagiu and J. Wright. *Multi-sided platforms*. Harvard Business School, 2011.

[52] C. Hajaj, J. P. Dickerson, A. Hassidim, T. Sandholm, and D. Sarne. Strategy-proof and efficient kidney exchange using a credit mechanism. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*, pages 921–928, 2015.

[53] C. Hajaj, N. Hazon, and D. Sarne. Improving comparison shopping agents' competence through selective price disclosure. *Electronic Commerce Research and Applications*, 14(6):563–581, 2015.

[54] C. Hajaj, N. Hazon, and D. Sarne. Enhancing comparison shopping agents through ordering and gradual information disclosure. *In Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-16)*, pages 1–19, 2016.

[55] C. Hajaj, N. Hazon, D. Sarne, and A. Elmalech. Search more, disclose less. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, 2013.

[56] C. Hajaj and D. Sarne. Strategic information platforms: selective disclosure and the price of free. In *Proceedings of the fifteenth ACM conference on Economics and Computation*, pages 839–856. ACM, 2014.

[57] M. Hajiaghayi, R. Kleinberg, M. Mahdian, and D. Parkes. Online auctions with re-usable goods. In *Proceedings of the sixth ACM Conference on Economics and Computation*, pages 165–174, 2005.

[58] T. K. Hartnett and T. Offerman. Efficiency in auctions with private and common values: An experimental study. *The American Economic Review*, 92(3):625–643, 2002.

[59] E. Horvitz, C. Kadie, T. Paek, and D. Hovel. Models of attention in computing and communication: from principles to applications. *Communications of the ACM*, 46(3):52–59, 2003.

[60] B. Hui and C. Boutilier. Who's asking for help?: a bayesian approach to intelligent assistance. In *Proceedings of IUI*, pages 186–193, 2006.

[61] M. C. Janssen and J. L. Moraga-González. Strategic pricing, consumer search and the number of firms. *The Review of Economic Studies*, 71(4):1089–1118, 2004.

[62] I. Jewitt and D. Li. Cheap-talk information disclosure in auctions. Technical report, Working paper, University of Oxford, 2012.

[63] J. P. Johnson and D. P. Myatt. On the simple economics of advertising, marketing, and product design. *The American Economic Review*, 96(3):756–784, 2006.

[64] A. Juda and D. Parkes. An options-based solution to the sequential auction problem. *Artificial Intelligence*, 173(7-8):876–899, 2009.

[65] D. Kahneman. A psychological point of view: Violations of rational rules as a diagnostic of mental processes (commentary on stanovich and west). *Behavioral and Brain Sciences*, 23:681–683, 2000.

[66] E. Kamar, Y. K. Gal, and B. J. Grosz. Modeling information exchange opportunities for effective human–computer teamwork. *Artificial Intelligence*, 195:528–550, 2013.

[67] J. O. Kephart and A. R. Greenwald. Shopbot economics. In *Game Theory and Decision Theory in Agent-Based Systems*, pages 119–158. Springer, 2002.

[68] J. O. Kephart, J. E. Hanson, and A. R. Greenwald. Dynamic pricing by software agents. *Computer Networks*, 32(6):731–752, 2000.

[69] G. Keren. Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 39(1):98–114, 1987.

[70] P. Klemperer. Auction theory: A guide to the literature. *Journal of economic surveys*, 13(3):227–286, 1999.

[71] P. Klemperer. Auctions: theory and practice. *Judgment in Managerial Decision Making*, 2004.

[72] A. Klos, E. U. Weber, and M. Weber. Investment decisions and time horizon: Risk perception and risk behavior in repeated gambles. *Management Science*, 51(12):1777–1790, 2005.

[73] S. Kraus. Human-agent decision-making: Combining theory and practice. In *TARK-2015*, 2015.

[74] V. Krishna. *Auction Theory*. Academic Press, 2002.

[75] M. G. Lagoudakis, E. Markakis, D. Kempe, P. Keskinocak, A. J. Kleywegt, S. Koenig, C. A. Tovey, A. Meyerson, and S. Jain. Auction-based multi-robot routing. In *Robotics: Science and Systems*, volume 5, page 343C350. Rome, Italy, 2005.

[76] K. P. Lai, S. C. Chong, H. B. Ismail, and D. Y. K. Tong. An explorative study of shopper-based salient e-servicescape attributes: A Means-End Chain approach. *International Journal of Information Management*, 34(4):517–532, 2014.

[77] R. Lavi and N. Nisan. Competitive analysis of incentive compatible on-line auctions. In *Proceedings of the 2nd ACM Conference on Electronic Commerce*, pages 233–241. ACM, 2000.

[78] P. Levy and D. Sarne. Intelligent advice provisioning for repeated interaction. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, pages 842–849, 2016.

[79] C. Lim, J. Bearden, and J. Smith. Sequential search with multiattribute options. *Decision Analysis*, 3(1):3–15, 2006.

[80] J. B. MacQueen. Optimal policies for a class of search and evaluation problems. *Management Science*, 10(4):746–759, 1964.

[81] P. M. Markopoulos and L. H. Ungar. Shopbots and pricebots in electronic service markets. In *Game Theory and Decision Theory in Agent-Based Systems*, pages 177–195. Springer, 2002.

[82] E. Maskin and J. Riley. Monopoly with incomplete information. *RAND Journal of Economics*, 15:171–196, 1984.

[83] P. Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12:380–391, 1981.

[84] P. Milgrom and R. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, 1982.

[85] G. Moscarini and L. Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.

[86] Y. Nahum, D. Sarne, S. Das, and O. Shehory. Two-sided search with experts. *In Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*, 29(3):364–401, 2015.

[87] M. Nermuth, G. Pasini, P. Pin, and S. Weidenholzer. The informational divide. *Games and Economic Behavior*, 78:21–30, 2013.

[88] C. Ng, D. C. Parkes, and M. Seltzer. Virtual worlds: Fast and strategyproof auctions for dynamic resource allocation. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pages 238–239. ACM, 2003.

[89] N. Nisan. Algorithms for selfish agents. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 1–15. Springer, 1999.

[90] G. Paolacci, J. Chandler, and P. G. Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision making*, 5(5):411–419, 2010.

[91] D. Parkes and J. Shneidman. Distributed implementations of vickrey-clarke-groves mechanisms. In *Proceedings of the Third International Conference on Autonomous Agents and Multiagent Systems (AAMAS-04)*, volume 1, pages 261–268, 2004.

[92] N. Peled, S. Kraus, et al. A study of computational and human strategies in revelation games. *In Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*, 29(1):73–97, 2015.

[93] M. Perry and P. Reny. On the failure of the linkage principle in multi-unit auctions. *Econometrica*, 67(4):895–900, 1999.

[94] M. Rabin. Psychology and economics. *Journal of Economic Literature*, pages 11–46, 1998.

[95] I. Rochlin, Y. Aumann, D. Sarne, and L. Golosman. Efficiency and fairness in team search with self-interested agents. *In Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-16)*, 30(3):526–552, 2016.

[96] I. Rochlin and D. Sarne. Utilizing costly coordination in multi-agent joint exploration. *Multiagent and Grid Systems*, 10(1):23–49, 2014.

[97] I. Rochlin and D. Sarne. Constraining information sharing to improve cooperative information gathering. *Journal Artificial Intelligence Research (JAIR)*, 54:437–469, 2015.

[98] I. Rochlin, D. Sarne, and M. Mash. Joint search with self-interested agents and the failure of cooperation enhancers. *Artificial Intelligence*, 214:45–65, 2014.

[99] I. Rochlin, D. Sarne, and G. Zussman. Sequential multi-agent exploration for a common goal. *Web Intelligence and Agent Systems*, 11(3):221–244, 2013.

[100] A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488, 2004.

[101] M. Rysman. The economics of two-sided markets. *The Journal of Economic Perspectives*, 23(3):125–143, 2009.

[102] D. Sarne, S. Alkoby, and E. David. On the choice of obtaining and disclosing the common value in auctions. *Artificial Intelligence*, 215:24–54, 2014.

[103] D. Sarne and B. J. Grosz. Sharing experiences to learn user characteristics in dynamic environments with sparse data. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-07)*, pages 202–209, 2007.

[104] D. Sarne, M. Hadad, and S. Kraus. Auction equilibrium strategies for task allocation in uncertain environments. In *Proceedings of the Eight International Workshop on Cooperative Information Agents (CIA)*, pages 271–285, 2004.

[105] D. Sarne and S. Kraus. The search for coalition formation in costly environments. In *Proceedings of the 7ht International Workshop on Cooperative Information Agents (CIA)*, pages 117–136. Springer, 2003.

[106] D. Sarne and S. Kraus. Solving the auction-based task allocation problem in an open environment. In *Proceedings of the national conference on artificial intelligence*, volume 20, pages 164–169, 2005.

[107] R. Schwartz and S. Kraus. Bidding mechanisms for data allocation in multi-agent environments. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 61–75. Springer, 1997.

[108] I. Sheena. *The Art of Choosing*. Twelve, March 2010.

[109] H. A. Simon. Theories of bounded rationality. *Decision and organization: A volume in honor of Jacob Marschak*, pages 161–176, 1972.

[110] L. Sless, N. Hazon, S. Kraus, and M. Wooldridge. Forming coalitions and facilitating relationships for completing tasks in social networks. In *Proceedings of the13th international conference on Autonomous agents and multi-agent systems(AAMAS-14)*, pages 261–268, 2014.

[111] P. Stone, R. Schapire, M. Littman, J. Csirik, and D. McAllester. Decision-theoretic bidding based on learned density models in simultaneous, interacting auctions. *Journal of Artificial Intelligence Research (JAIR)*, 19(1):209–242, Sept. 2003.

[112] M. Tennenholtz. Tractable combinatorial auctions and b-matching. *Artificial Intelligence*, 140(1-2):231–243, 2002.

[113] R. H. Thaler and C. R. Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, 2008.

[114] I. Vetsikas and N. Jennings. Bidding strategies for realistic multi-unit sealed-bid auctions. *In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-10)*, 21(2):265–291, 2010.

[115] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

[116] R. Waldeck. Search and price competition. *Journal of Economic Behavior & Organization*, 66(2):347–357, 2008.

[117] Y. Wan and G. Peng. What's next for shopbots? *IEEE Computer*, 43(5):20–26, 2010.

[118] E. G. Weyl. A price theory of multi-sided platforms. *The American Economic Review*, 100(4):1642–1672, 2010.

[119] D. D. Wiegmann, K. L. Weinersmith, and S. M. Seubert. Multi-attribute mate choice decisions and uncertainty in the decision process: a generalized sequential search strategy. *Journal of mathematical biology*, 60(4):543–572, 2010.

[120] K. E. Wilson, R. Szechtman, and M. P. Atkinson. A sequential perspective on searching for static targets. *European Journal of Operational Research*, 215(1):218–226, 2011.

[121] R. Wilson. Nonlinear pricing. *Oxford University Press*, 1993.

[122] P. R. Wurman, M. P. Wellman, and W. E. Walsh. A parametrization of the auction design space. *Games and economic behavior*, 35(1):304–338, 2001.

[123] M. Yakout, A. K. Elmagarmid, J. Neville, M. Ouzzani, and I. F. Ilyas. Guided data repair. *Proceedings of the VLDB Endowment*, 4(5):279–289, 2011.

[124] M. Yokoo, Y. Sakurai, and S. Matsubara. Robust combinatorial auction protocol against false-name bids. *Artificial Intelligence*, 130(2):167–181, 2001.

ספקי מידע אסטרטגיים וחשיפת מידע סלקטיבית במערכות מרובות סוכנים

חיבור לשם קבלת התואר "דוקטור לפילוסופיה"

מאת:
שני אלקובי

המחלקה למדעי המחשב

הוגש לסנאט של אוניברסיטת בר אילן

רמת גן　　　　　　　　　　　　　　אב, תשע"ז

עבודה זו נעשתה בהדרכתו של פרופ׳ דוד סרנה
מן המחלקה למדעי המחשב של אוניברסיטת בר אילן.

תוכן העניינים:

תקציר

המחקר המדווח בתזה זו מספק תובנות עמוקות לגבי תפקיד המידע והשימוש האפקטיבי בו במערכות מרובות סוכנים. בפרט, אני מתמקדת בשלוש סביבות מרכזיות: מכרזים, חיפוש כלכלי וממשק אדם ומכונה, בהן אני ממחישה בצורה משמעותית כיצד משפיע על המערכת המידע שסופק. שימוש אפקטיבי במידע הינו תחום מחקר פופולרי כיום, בעיקר הודות לתפקיד המפתח שמשחק המידע בחיי היומיום של כולנו. הוכחה לכך ניתן למצוא במספר ההולך וגדל של חברות מסחריות המשקיעות נתח משמעותי מהמשאבים שלהן במחקר העוסק בטכנולוגיות מידע. לדוגמה, בשנת 2014, רכשה גוגל חברת סטארט-אפ העוסקת באספקת מידע, בסכום של 650 מיליון דולר, מה שהופך אותה לאחת מהרכישות הגדולות ביותר שנעשו עד כה. יתר על כן, לפחות ל 30% מתוך רשימת 100 החברות הציבוריות הגדולות ביותר על פי מגזין פורבס יש קשר הדוק לתחומים של אספקת מידע. למעשה, על פי האקונומיסט, המשאב היקר ביותר בעולם כיום הוא לא נפט, אלא מידע.

בספרות הקיימת ישנה התמקדות רבה בתפקיד המידע בסביבות שונות. כאשר מתמודדים עם אספקת מידע, שאלות רבות צריכות להיענות, למשל: מהו הערך של המידע? האם יש לחשוף את כל המידע או רק חלק ממנו? באיזו דרך יש להציג את המידע? האם המידע ניתן בחינם או תמורת תשלום כלשהו? האם גילוי המידע יהיה אנונימי? כיצד יש להשתמש במידע כדי למקסם את הרווח הצפוי מהמערכת? ועוד מספר רב של שאלות. בתזה זו אני חוקרת מספר דרכים לאספקת מידע יעיל עבור סוכנים במערכות מרובות סוכנים כאשר ספקי המידע הינם סוכנים רציונליים המעוניינים במקסום הרווח האישי שלהם.

עבור כל אחד מהתחומים שהוזכרו, ראשית אני מציגה מהו תפקיד המידע במערכת, לאחר מכן, אני מראה כיצד גילוי המידע יכול להשפיע על המשתתפים השונים במערכת (תוך שימוש בניתוח תיאורטי כאשר הדבר אפשרי (כלומר כאשר מתמודדים עם גורמים רציונליים)) ולבסוף מספקת דרכים על מנת להגדיל את הרווח הצפוי של המשתתפים השונים תוך שימוש בגילוי מידע אפקטיבי.

התחום הראשון שבו אני עוסקת הינו מכרזים. בתחום זה, המוצג בפרק 3, אני בוחנת כיצד סכמות שונות של גילוי מידע משפיעות על ההתנהגות ועל הרווח הצפוי של המשתתפים השונים במכרז (ספק המידע, מנהל המכרז והקונים), ובנוסף על התועלת החברתית. בתחום זה אני מראה כי ישנם מקרים שבהם על ידי חשיפה חלקית חינמית של המידע, ספק המידע יכול להוביל את השוק לשיווי משקל טוב יותר עבורו, כזה שמעניק לו רווח צפוי גבוה יותר.

במקרים רבים, לצד מכירת המידע, השחקן המחזיק במידע יכול לבצע פעולות נוספות. פעולות אלו מאפשרות לספק המידע לשנות את האמונות של משתתפים אחרים במכרז בנוגע לערך האמיתי של הפריט המוצע למכירה. מנהל המכרז, אפילו שאינו מחזיק במידע, יכול גם הוא להשפיע על האמונות של הקונים לגבי הערך האמיתי של הפריט המוצע למכירה. בפרק 4 אני דנה בשתי פעולות אפשריות נוספות (ניתן גם להתייחס אליהן כאל מניפולציות). המניפולציה הראשונה שבה אני דנה היא זו שמתבצעת על ידי עורך המכרז. עורך המכרז יכול לפגוע ביכולת של ספק המידע להבחין בין כמה מן הערכים האפשריים. כלומר, עורך המכרז למעשה גורם למידע שנחשף אליו להיות פחות מדויק (בהשוואה למידע שהיה מקבל ללא שימוש במניפולציה זו). המניפולציה השנייה שבה אני דנה, היא זו שבה ספק המידע חושף מידע בחינם בצורה אנונימית כך שעורך המכרז והקונים אינם יודעים שהוא מקור המידע. הבחירה באנונימיות, מאפשרת לספק המידע למנוע את התגובה האסטרטגית של המשתתפים האחרים (בשל העובדה שהם אינם מודעים להתנהגותו האסטרטגית). במקרה כזה, האסטרטגיה הדומיננטית של ספק המידע היא לשלול תת-קבוצה של ערכים כך שהרווח הצפוי שלו מהמכרז יהיה מקסימלי. למרות שלכאורה מדובר על החלטה פשוטה, בפועל היא כרוכה במורכבות חישובית גדולה. על מנת להעניק לספק המידע כלי מעשי שיסייע לו לבחור את הערכים אותם הוא מעוניין לשלול, אני מציגה שתי שיטות היוריסטיות לסידור רצף הפתרונות האפשריים, כך שאלו המניבים רווח גבוה יותר יופיעו בהסתברות גבוהה מוקדם יותר ברצף. שתי היוריסטיקות אלו נבחנו

ונמצאו יעילות ביותר ניסויית. בנוסף, אני מראה שחשיפת מידע על ידי ספק המידע לכל השחקנים ולא רק לעורך המכרז יכולה לעיתים לשפר את הרווח הצפוי של ספק המידע (למרות שעורך המכרז הוא המשתתף היחיד המורשה לרכוש את המידע).

התחום השני שבו אני דנה בתזה זו הוא חיפוש כלכלי (ובפרט חיפוש כלכלי חד צדדי). המכנה המשותף לרוב המחקרים העוסקים בחיפוש כלכלי הוא המודל הבסיסי הכולל סט הזדמנויות המוצעות לסוכן, מהן הוא רשאי לבחור אחת בלבד. על מנת למצוא את ההזדמנות המיטבית עבורו, מבצע סוכן זה חיפוש בין ההזדמנויות המוצעות לו. חשוב לציין שהסוכן המחפש מעוניין למקסם לא רק את הרווח הצפוי לו מההזדמנות שאותה יבחר בסופו של דבר, אלא את סך הרווח הכולל, המושפע גם מעלות תהליך החיפוש. בתזה זו, אני עוסקת באספקת מידע בחיפוש כלכלי חד צדדי. דוגמא פופולרית לחיפוש חד צדדי היא המקרה של חיפוש עבודה, במקרה זה, מחפש העבודה מעוניין למצוא את העבודה המתאימה ביותר עבורו אך עליו לקחת בחשבון בנוסף לרווח הצפוי לו ממציאת העבודה המתאימה ביותר עבורו, את העלות (למשל הזמן) שהוא משקיע בחיפוש. בפרק 5 אני בוחנת מקרים בסביבה של חיפוש כלכלי בהם ספק המידע מעוניין לחשוף חלק מהמידע בחינם.

בסביבה של חיפוש חד צדדי, המידע לגבי איכות ההזדמנות הוא בעל חשיבות רבה לסוכן המבצע את החיפוש. בתזה זו אני מראה כי מתן אפשרות לספק מידע להשתמש בגילוי מידע חלקי בחינם יכול להוביל את הסוכן המחפש לרכוש את שירותיו של ספק המידע תמורת תשלום גם במקרים בהם הוא לא התכוון לעשות זאת מלכתחילה (כלומר, כאשר לא נחשף מידע חינמי). בנוסף, אני מראה ששימוש בחשיפה חלקית של המידע בחינם יכול להוביל לעיתים קרובות לרווח גבוה יותר עבור ספק המידע (ולפעמים אפילו עבור הסוכן המחפש). יתר על כן, אני מצליחה להציג מבנה ייחודי של שיווי משקל המתקיים במקרה זה.

תופעה מעניינת (אפשר אפילו לומר טבעית) שעלולה להתרחש במקרים כאלה שבהם ספק המידע מציע מידע חינמי, היא ניצול לרעה של ספק המידע. ניצול לרעה יכול לקרות כאשר סוכנים נוספים, אשר לא היו מעוניינים לשלם עבור המידע, יהיו מעוניינים כעת להשתמש במידע בשל היותו בחינם. כתוצאה מכך, הרווח הצפוי של ספק המידע עשוי להיפגע (עקב הצורך ביצירת מידע רב הניתן בחינם). במהלך התזה אני דנה בתרחיש זה ומצליחה להראות כי במודל ספציפי זה, הרווח הצפוי של ספק המידע אינו נפגע אפילו במקרים שבהם יש מספר רב של סוכנים המנצלים את המידע החינמי.

התחום האחרון שבו אני דנה הינו השפעת מידע על אנשים. במצבים רבים בחיים האמיתיים, הסוכנים המעוניינים במידע הם בני אדם (כלומר סוכנים שאינם רציונליים לחלוטין). לכן, בפרק 6 אני חוקרת את המקרה שבו קונייני המידע הם אנשים. התוצאות שהושגו עבור המקרה שבו המידע נמכר לאנשים מושוות לשיוויי המשקל במקרה שבו קונייני המידע הם סוכנים רציונאליים לחלוטין, על מנת להדגיש את ההבדלים ביניהם. אני מראה שלמרות העובדה שכאשר מתמודדים עם סוכנים רציונליים, ספק המידע אינו יכול להפיק תועלת משימוש בחשיפת מידע חלקי בחינם, זהו לא המקרה כאשר מדובר באנשים. אחת התוצאות המעניינות והמפתיעות שאני מציגה בתזה זו מראה שלמרות שאנשים מתקשים בביצוע החישוב המדויק של הערך הנכון של המידע, זו לא הסיבה העיקרית להבדל בין התוצאות התיאורטיות לבין המתרחש בפועל. הסיבה העיקרית לעלייה ברווח הצפוי של ספק המידע כאשר הוא נמצא באינטראקציה עם אנשים היא שאנשים לא לוקחים בחשבון את העובדה שספק המידע הינו שחקן אסטרטגי.

המחקר המתואר בתזה זו מבוסס על ניתוח תיאורטי ועל ניסויים אמפיריים מקוונים. הניתוח התיאורטי מתבצע על ידי שימוש במושגים מתורת המשחקים, תורת המכרזים ותורת החיפוש, בעוד שהניסויים המקוונים מבוססים על פלטפורמת ניסויים של אמזון, פלטפורמת מיקור המונים ידועה.

לסיכום, אני מאמינה שאספקה יעילה של מידע היא כלי חזק מאוד וניתן להשתמש בו בסביבות רבות בחיי היומיום, ובמיוחד במערכות מרובות סוכנים, על מנת להשפיע על ההתנהגות של הסוכנים המשתתפים ולהוביל אותם לתוצאות יעילות יותר. אני מאמינה שאפשר למנף את התוצאות שהושגו עד כה על מנת ליצור תורה סדורה של אספקת מידע בצורה יעילה שתהיה מאוד שימושית ומעשית, נוכח העובדה שהמידע הוא מרכיב חיוני והכרחי בחיי היומיום שלנו.