

Making Money from What You Know - How to Sell Information?

Shani Alkoby,¹ Zihe Wang,² David Sarne,³ Pingzhong Tang⁴

¹The University of Texas at Austin, Austin, Texas, USA

²Shanghai University of Finance and Economics, Shanghai, China

³Bar-Ilan University, Ramat Gan, Israel ⁴Tsinghua University, Beijing, China

shani.alkoby@gmail.com, wzh5858588@163.com, sarned@cs.biu.ac.il, kenshinping@gmail.com

Abstract

Information plays a key role in many decision situations. The rapid advancement in communication technologies makes information providers more accessible, and various information providing platforms can be found nowadays, most of which are strategic in the sense that their goal is to maximize the providers' expected profit. In this paper, we consider the common problem of a strategic information provider offering prospective buyers information which can disambiguate uncertainties the buyers have, which can be valuable for their decision making. Unlike prior work, we do not limit the information provider's strategy to price setting but rather enable her flexibility over the way information is sold, specifically enabling querying about specific outcomes and the elimination of a subset of non-true world states alongside the traditional approach of disclosing the true world state. We prove that for the case where the buyer is self-interested (and the information provider does not know the true world state beforehand) all three methods (i.e., disclosing the true world-state value, offering to check a specific value, and eliminating a random value) are equivalent, yielding the same expected profit to the information provider. For the case where buyers are human subjects, using an extensive set of experiments we show that the methods result in substantially different outcomes. Furthermore, using standard machine learning techniques the information provider can rather accurately predict the performance of the different methods for new problem settings, hence substantially increase profit.

1 Introduction

1.1 Discussion

One interesting trend reflected in the Fortune 500 lists in recent years is the increase in the number of companies whose primary business is producing and selling information rather than traditional tangible goods. The types and natures of the information offered and the business models used for selling it greatly vary. Yet the goal of all the firms offering it is similar - maximizing their profit. Hence there is a growing interest in finding ways for improving the effectiveness of information offerings and their pricing.

In this paper we consider the problem of a strategic information provider who can obtain (or calculate or predict)

the true state of the world, out of several possible states, and provide this information to prospective information buyers. This setting is common in real life. For example, the information can relate to the true value of a car (e.g., by Kelley Blue Book or Carfax), the true worth of a company the buyer is interested in taking over, the exact weather for one's wedding day and the true amount of oil buried in a given land a firm owns. Common to all these examples, that the true state of the world is a priori uncertain, and by disambiguating the uncertainty associated with it the information buyer can make better decisions hence increase her profit.

The analysis provided considers and compares three information selling methods for the above archetypal setting. The first is the traditional method, according to which the exact state of the world is delivered for a fee. The second enables the information buyer to query about any specific potential world state for a fee (different for each world state, and changing as the process progresses). The third offers the elimination of one world state randomly picked from the set of non-true world states, for a fee (which changes as the process progresses). If considering the above mentioned oil example, assuming there are several a priori estimates for the amount of oil buried under the ground, an expert can use the first method and set a price for selling the exact information. Alternatively, the seller can offer the buyer to ask specifically whether a certain value is the true one or not. Finally, the seller can also offer to eliminate one non-true option chosen randomly and by doing so to decrease the number of states the buyer should further consider.

The first part of the paper provides a game-theoretic based analysis of information selling using the three methods. We show that all three methods are equivalent in the sense that the information provider's expected profit in equilibrium is the same regardless of the method used. In the second part, we report the result of extensive experiments with the three methods when selling the information to people rather than fully rational agents. Here, we show that the three methods result in quite different outcomes in terms of the information provider's profit. Furthermore, using a standard machine learning technique (decision trees) we show that the information provider can benefit much from selectively choosing the method to be used (and the corresponding pricing of the information offered) based on the problem setting. These results are of great importance both to information

providers who want to improve their information offerings and for market designers and regulators who can often constrain the way information is being sold in a way that provides a decent tradeoff between information providers' profits and social welfare.

1.2 Related work

Being a well-researched topic, information providing has been studied from various different points of view. A considerable number of studies have been studying ways for increasing the expected profit of the information buyers (Yakout et al. 2011; Alkoby, Sarne, and David 2014; Horvitz et al. 2003; Hui and Boutilier 2006; Alkoby et al. 2018; Hajaj, Hazon, and Sarne 2015), whereas others have tried maximizing the overall social welfare (Sarne, Alkoby, and David 2014). Still, most works have considered, similar to this paper, the problem of maximizing the expected profit of the information provider herself, offering various methods. This included primarily mechanisms for the information providers for how they should price their services (Moscarini and Smith 2003; Azoulay-Schwartz and Kraus 2004; Cheng and Koehler 2003; Lai et al. 2014), as well as methods such as using partially free information disclosure (Alkoby, Sarne, and Milchtaich 2017; Ganuza and Penalva 2010; Alkoby, Sarne, and Das 2015; Dufwenberg and Gneezy 2002; Esó and Szentes 2007; Alkoby and Sarne 2015), or controlling the order of presenting information to prospective information buyers (Hajaj, Hazon, and Sarne 2017). Specifically, Hajaj et al. (Hajaj and Sarne 2017) showed that information platforms should take the subset of opportunities to be included in their listings as a decision variable, alongside the fees set for the service in their expected-profit "maximizing" optimization problem. These were extended also for cases where the prospective buyers are humans (Hotz and Xiao 2013; Azaria et al. 2015). Alas, all those work limited themselves to a direct sale of the information, i.e., if purchasing the information the buyer gains certainty regarding the true state of the world.

The three most relevant works to our research are our previous work (Alkoby and Sarne 2017) and those of (Daskalakis, Papadimitriou, and Tzamos 2016) and (Bergemann, Bonatti, and Smolin 2018). In our previous work, the information seller attempts to increase her expected profit by using partial free information disclosure. There also, we consider a direct sale of the true world state. Daskalakis et al. assumes buyers to be completely rational and tries to find the optimum co-design of signaling in an auction. This work is a part of the signaling literature, in which a portion of the information is being disclosed for free in the form of a signal in order to influence the buyer's decision, and hence the seller's expected profit (Bro Miltersen and Sheffet 2012; Emek et al. 2011). Bergemann et al. assume buyers rationality in addition for assuming heterogeneity in their beliefs and needs. Their work gives a great importance to finding the properties that any revenue-maximizing menu of experiments must satisfy whereas in our case, the emphasis is not on how to build a menu of offers suited for each one of the buyers, but rather given a human buyer, how should an in-

formation provider sell the information she owns in a way that will maximize her profit.

To the best of our knowledge, an empirical investigation of methods different than providing the exact state of the world when facing human buyers has not been carried out to date.

2 Model and Methods

We consider a standard model of a self-interested information seller (*IS*) and a prospective information consumer (*IC*). The *IC* faces a simple binary decision problem with two available actions (e.g., buy or not buy, drill or not drill, carry out wedding ceremony indoor or in the garden area). WLOG we assume that the profit of the *IC* from one of the alternatives (denoted "opt out" onwards) is v_0 .¹ The profit from the second alternative (denoted "exploit" onwards) depends on some future world state which is a priori uncertain. We use $V = \{v_1, \dots, v_k\}$ to denote the *IC*'s profit in case of choosing the second alternative given the different world states, such that v_i is the profit in case the true world state turns to be i . The corresponding a priori probability of each world state i (and consequently the corresponding value $v_i \in V$) is captured by the function p_i ($\sum p_i = 1$).

Both players are symmetric in the sense that they are acquainted with the set V and the underlying probability function p_i . Unlike the *IC*, the *IS* can check for each potential world state if it is indeed the true world state at the time outcomes are set and can offer such service to the *IC*. Meaning that the information provider does not initially have the information but rather produce it on the spot. This is common whenever the production of information is costly (hence no point in producing it ahead of time) or when obtaining the information requires some complementary data from the buyer herself (e.g., estimating the true amount of oil buried in a given land requires receiving all the data produced in exploratory drills carried out by the buyer).

The paper considers and studies three methods according to which the *IS*'s service can be offered to the *IC*:

- *Full Information* - with this method, which is traditionally the one considered in literature for selling information in such settings (Alkoby and Sarne 2017), upon request the *IS* obtains the true value v and delivers it to the *IC* for a fee.
- *Options Menu* - with this method the information is being sold in a sequential process, according to which at each stage the seller publishes a menu where each item i in the menu specifies the price for knowing, at that stage, whether or not value v_i is the true one. If the buyer chooses to query for one of the values in the menu then the seller reveals whether this is indeed the true value. If it is the true value then the process terminates. Otherwise, the seller sets a new menu for querying any of the remaining set of values and so on. At any step of the process

¹Alternatively, assume the *IC* has an opportunity available to her, where the possible alternatives are to exploit that opportunity or opt not to exploit it, in which case her profit is v_0 .

the buyer may opt to terminate the process and act optimally based on the information accumulated so far along the process.

- *Random Elimination* - here also the information is being sold in a sequential process. At each stage the *IS* publishes a price for randomly eliminating one of the non-true values. The *IC* can either accept or opt to terminate the process and act optimally based on the information accumulated so far along the process. If accepting, then the process continues in the same manner with a new price set by the *IS* for eliminating a non-true value out of the remaining ones, and so on.

3 Fully Rational Information Consumers

Naturally, the price set for the service should be equal to the value of the information offered to the *IC* in any of the methods. Setting a greater price will result in not purchasing the information, and setting a price lower than the value of the information is dominated by the latter. Theorem 1 suggests that despite the inherent differences in the way information is being sold in each one of the above mentioned three methods, the methods are actually equivalent in the sense that the *IS* makes the same expected profit with all three, whenever the *IC* is fully rational.

Theorem 1 *With a fully rational IC, the expected profit of the IS when using any of the three methods is identical, and equals:*

$$\sum_{i=1}^k p_i \max(v_i, v_0) - \max\left(\sum_{i=1}^k p_i v_i, v_0\right) \quad (1)$$

Similarly the expected profit of the IC is the a priori expected value of choosing to exploit, and equals:

$$\max\left(\sum_{i=1}^k p_i v_i, v_0\right) \quad (2)$$

Proof. In the absence of any information from the *IS*, the *IC* will choose *exploit* only if the expected value of this alternative is greater than the fallback value v_0 . Hence, if not receiving any information from the *IS*, the *IC*'s expected profit is $\max\left(\sum_{i=1}^k p_i v_i, v_0\right)$. With full information, upon receiving the true value v the *IC* will choose *exploit* if $v > v_0$ and *opt out* otherwise. Therefore if receiving the information her expected profit is $\sum_{i=1}^k p_i \max(v_i, v_0)$, and consequently the price the *IS* will charge for the information is according to 1, leaving the *IC* with profit (2).

In order to prove that the *IS*'s profit is similar with the two other methods, we first introduce Lemma 2, showing that both with Options Menu and Random Elimination, the posterior probability of any of the remaining (non-eliminated) values is the same and depends only on the remaining set of value itself.

Lemma 2 *After any sequence of operations, if the remaining possible set of values is S , the probability of $v_i \in S$ being the true value is $\frac{p_i}{p(S)}$, where $p(S)$ is the sum of the probabilities of elements in set S , i.e. $p(S) = \sum_{v_j \in S} p_j$.*

Proof. The proof applies to any sequence of operations (i.e., both checking a specific value and realizing it is not the true world state and randomly removing non-true value). We prove by induction, thus, all we need is to prove that the lemma holds given that a single operation was made.

Let the current set of possible values be $S = \{v_1, v_2, \dots, v_l\}$. We begin with the Options Menu. Assume the value v_i is picked by the *IC* and eliminated, i.e., the remaining set is $S_i = S - v_i = \{v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{|S|}\}$. Since the elimination of v_i does not reveal new information on any of the values in S_i , the posterior probability that a value $v_j \in S_i$ is the true one, denoted $\Pr(v_j|S_i)$, is given by:

$$\Pr(v_j|S_i) = \frac{p_j}{p(S_i)}$$

Now consider the Random Elimination when v_i is eliminated. Since v_i was randomly picked, the posterior probability that a value $v_j \in S_i$ is the true one is given by:

$$\begin{aligned} \Pr(v_j|S_i) &= \frac{\Pr(S_i|v_j)p_j}{\sum_{v_w \in S_i} \Pr(S_i|v_w)p_w} \\ &= \frac{\frac{1}{l-1}p_j/p(S)}{\sum_{v_w \in S_i} \frac{1}{l-1}p_w/p(S)} = \frac{p_j}{p(S_i)} \end{aligned}$$

Where $\Pr(S_i|v_j)$ is the probability that v_i was picked for elimination, i.e., we ended up with S_i , given that v_j is the true value. Note that $\Pr(S_i|v_j) = 1/(l-1)$ because any of the values other than v_j has an equal chance of being selected for elimination.

Therefore the lemma holds after any single operation. Applying this argument over and over again, we can prove that the lemma holds after any sequence of operations. \square

The fact that the posterior probability of any of the remaining values depends only on the remaining set itself simplifies the problem as one does not need to consider the entire history of operations leading to a given set.

We begin with the Options Menu method and prove the theorem's claim for this case by induction. Assume the set of remaining non-eliminated values S includes two elements. Here, the price that will be set by the *IS* is the same for both values (and equal to the price of providing the true outcome out of the two), as asking about one outcome is equivalent to asking about the true outcome out of the two - having the information for one outcome necessarily reveals the nature of the other. Now assume that for $|S| < i$ the theorem holds (meaning that the *IS* is indifferent between offering to sell the true outcome and offering to provide information about one outcome the *IC* will choose, and so is the *IC*). We will use the following assisting notations: B_S^{IS} will denote the expected profit of the *IS* if acting optimally when the set of values is S ; $B_S^{IS}(om)$ will denote the expected profit of the *IS* onwards if offering to check one of the values (Options Menu) and the set of values is S ; $B_S^{IS}(fi)$ will be the expected profit of the *IS* if offering the exact value and the set of values is S ; and B_S^{IC} will denote the expected profit of the *IC* if the current set of values is S . For simplicity, we use x^+ to denote $\max\{x, v_0\}$ for the rest of the paper.

According to the proof assumption, for any $|S| < i$ we know: $B_S^{IS} = B_S^{IS}(om) = B_S^{IS}(fi)$. The expected profit of the IS , $B_S^{IS}(fi)$, is given by:

$$B_S^{IS}(fi) = \sum_{v_j \in S} \frac{p_j}{p(S)} \cdot v_j^+ - E[S]^+ \quad (3)$$

This is because the IS can charge exactly the amount the information is worth to the IC where $E[S] = \sum_{i=1}^{|S|} p_i v_i^+ / p(S)$. This is because with probability $\frac{p_i}{p(S)}$ v_i turns to be the true value and the profit of the IC is $\max\{v_i, v_0\}$ and with the complementing probability it is not the true value and the expected profit of the IC is $B_{S-v_i}^{IS}$. Without the information (i.e., if terminating the process) the profit is $E[S]^+ = \max\{E[S], v_0\}$. Now we need to prove that this holds also for $|S| = i$. The price the IS should set for checking a value $v_i \in S$ is $\frac{p_i}{p(S)} v_i^+ + (1 - \frac{p_i}{p(S)}) B_{S-v_i}^{IC} - \max\{E[S], v_0\}$. The IC 's expected profit remains $\max\{E[S], v_0\}$ either way (because the IS takes all the surplus). Therefore, if offered to buy a value then the IC is indifferent between checking any of the values. Hence the IS profit from having the IC asking to check on value $v_i \in S$ is given by:

$$\begin{aligned} & \frac{p_i}{p(S)} v_i^+ + (1 - \frac{p_i}{p(S)}) B_{S-v_i}^{IC} - E[S]^+ + (1 - \frac{p_i}{p(S)}) B_{S-v_i}^{IS} \\ &= \frac{p_i}{p(S)} v_i^+ + (1 - \frac{p_i}{p(S)}) E[S - v_i]^+ - E[S]^+ \\ &+ (1 - \frac{p_i}{p(S)}) (\sum_{v_j \in S-v_i} \frac{p_j}{p(S) - p_i} \cdot v_j^+ - E[S - v_i]^+) \\ &= \sum_{v_j \in S} \frac{p_j}{p(S)} \cdot v_j^+ - E[S]^+ \end{aligned}$$

which is equal to (1) whenever $S=V$.

Next, we prove for Random Elimination. Here, again, we prove by induction. For $|S| = 2$, randomly removing one value among those that are not the true world state and providing the true one is the same as revealing the true value. Now assume that for $|S| < i$ the theorem holds (meaning that the IS is indifferent between offering to sell the true value and offering to remove a random value) and so is the IC . Once again we use the assisting notations: B_S^{IS} , $B_S^{IS}(fi)$, and B_S^{IC} . $B_S^{IS}(re)$ will denote the expected profit of the IS onwards if offering to eliminate randomly one outcome when the set of remaining outcomes is S . According to the proof assumption, for any $|S| < i$ we know: $B_S^{IS} = B_S^{IS}(re) = B_S^{IS}(fi)$. Now we need to prove that this holds also for $|S| = i$. According to the assumption, for any $|S| < i$ we know that: $B_S^{IC} = E[S]^+$.

The worth to the IC of removing one value is given by:

$$\begin{aligned} & \sum_{v_j \in S} (\frac{p_j}{p(S)} \cdot \sum_{v_w \in S-v_j} \frac{1}{i-1} B_{S-v_w}^{IC}) - E[S]^+ \\ &= \sum_{v_j \in S} (\frac{p_j}{p(S)} \sum_{v_w \in S-v_j} \frac{1}{i-1} E[S - v_w]^+) - E[S]^+ \quad (4) \end{aligned}$$

In the above calculation we iterate over all possible values v_j considering for each the case it is the true value. If v_j is the true value then one of the other values $v_w \in S - v_j$ will be removed with probability $\frac{1}{i-1}$. The IC will continue with a reduced set $S - v_w$ and her expected profit onward will be $B_{S-v_w}^{IC}$. The expected further profit from one random elimination is thus given by $\sum_{v_j \in S} \frac{p_j}{p(S)} \cdot \sum_{v_w \in S-v_j} \frac{1}{i-1} B_{S-v_w}^{IS}$.

Therefore, the IS 's profit from removing a random value when current values are those in S , $B_S^{IS}(re)$, is:

$$\begin{aligned} & \sum_{v_j \in S} (\frac{p_j}{p(S)} \cdot \sum_{v_w \in S-v_j} \frac{1}{i-1} E[S - v_w]^+) - E[S]^+ \quad (5) \\ &+ \sum_{v_j \in S} (\frac{p_j}{p(S)} \cdot \sum_{v_w \in S-v_j} \frac{1}{i-1} B_{S-v_w}^{IS}) \end{aligned}$$

and notice that according to the assumption, $B_{S-v_w}^{IS}$ is

$$\sum_{v_z \in S-v_w} \frac{p_z}{p(S) - p_w} \cdot v_z^+ - E[S - v_w]^+$$

Therefore substituting it in (5), $B_S^{IS}(re)$ becomes:

$$\begin{aligned} & \sum_{v_j \in S} \frac{p_j}{p(S)} \sum_{v_w \in S-v_j} \frac{1}{i-1} \sum_{v_z \in S-v_w} \frac{p_z}{p(S) - p_w} v_z^+ - E[S]^+ \\ &= \sum_{v_w \in S} \sum_{v_z \in S-v_w} \sum_{v_j \in S-v_w} \frac{p_z}{p(S)} \frac{1}{i-1} \frac{p_j}{p(S) - p_w} v_z^+ - E[S]^+ \\ &= \sum_{v_w \in S} \sum_{v_z \in S-v_w} \frac{p_z}{p(S)} \frac{1}{i-1} v_z^+ - E[S]^+ \\ &= \sum_{v_z \in S} \sum_{v_w \in S-v_z} \frac{p_z}{p(S)} \frac{1}{i-1} v_z^+ - E[S]^+ \\ &= \sum_{v_z \in S} \frac{p_z}{p(S)} v_z^+ - E[S]^+ \end{aligned}$$

which once again equals (1) when $S = V$. \square

4 Bounded Rational Information Consumers

While the above analysis holds when both agents are fully rational and not limited computationally, in reality most IC 's are human and are likely to be bounded rational, make mistakes, get bored and have many other imperfections (Rabin 1998; Kahneman 2000; Azaria et al. 2015; Buntain, Azaria, and Kraus 2014; Wang and Tang 2015; Hajaj, Hazon, and Sarne 2017). Therefore, despite the equivalence of profit proved in the former section, we hypothesize that not only the three methods for selling information result in different expected profit, but also that through intelligent selection of the method to be used and the price to be set for the information, the IS can substantially improve her profit, whenever selling the information to people.

In order to test the above underlying hypothesis, we use an augmented version of the experimental framework introduced in our previous work (Alkoby and Sarne 2017), called "What's In The Box?". It is based on a multi-round game in which in every round the user is being introduced to a box containing a prize (in game points). The value of the prize is a priori unknown to the user. Instead, she is provided with

a set of k possible values, both positive and negative, where one of which is the true value of the prize. The user needs to ultimately decide whether to open the box, hence collect the prize in it, and the dilemma arises since the value might be negative. Prior to making her decision, the user is offered some help in the form of some additional information by the system for a fee. A more detailed description of the framework can be found in (Alkoby and Sarne 2017), including various discussions and justifications for its design choices.

In our version of the game we had three different types of help offerings, replicating the three information selling methods analyzed in this paper. In the Full Information game variant the *IS* offers to reveal the exact value of the prize in the box, for a fee; in the Options Menu the *IS* allows the user to query regarding a specific value from the set of possible prize value; and in the Random Elimination game variant, the *IS* offers to randomly eliminate one of the false values in the set of potential values, for a fee.

The above infrastructure was used to collect data about people’s information purchase decisions with the different information providing methods for different requested fees. For this purpose we pre-generated a set of 250 core problem settings. Each core problem had $k \in \{3, 4, 5, 6, 7\}$ possible prize values (50 problems for every k value), where values were integers randomly picked within the range $[-50, 50]$. While the range of possible fees the *IS* can request is practically infinite, we used six price levels for each of the 250 core problem settings, as follows. We use the *Value of Information* concept, VoI , to denote an agent’s marginal benefit from having the information compared to not having it. Obviously different ways of information providing produce different VoI to the *IC*. Thus, for each tested information providing method, we first calculated the VoI according to either equation (1) (for the Full Information method), (4) (for the Options Menu method), or (5) (for the Random Elimination method). We then set fees to be $0.2 \cdot VoI$, $0.5 \cdot VoI$, $0.8 \cdot VoI$, $1.2 \cdot VoI$, $1.5 \cdot VoI$, and $1.8 \cdot VoI$, enabling a decent coverage of range and sufficient granularity of price segments.

Data was collected by having subjects playing “What’s In The Box?”. Subjects were recruited and interacted through Amazon Mechanical Turk (AMT) which has proven to be a well established method for data collection in tasks which require human intelligence (Paolacci, Chandler, and Ipeirotis 2010). As in the original “What’s In The Box?” game we used a “between subjects” design in order to prevent any carryover effect. A show-up fee was given to the users. Additionally, for every 5 game points a user earned, she received a bonus of one cent. Subjects received thorough instructions explaining the game rules, participated in at least two practice game rounds and had to pass a short quiz making sure they fully understand the task before allowed to proceed to the actual game rounds. Then each subject played 20 game rounds, each introducing a new information selling setting using a randomly drawn instance of a randomly drawn core setting (where instances of the same core settings differ in the price set) using the same information selling method, i.e., each subject experienced only with one of the three information selling methods. The instructions clearly indicated that

there is no correlation whatsoever between the settings used in the different rounds.

Overall we had 450 subjects playing the game such that each information providing method was used with 150 of them. Since each experienced with 20 instances, we had a total of 3000 information purchasing (or not purchasing) decisions for a total of 1500 setting instances (each differing in the core setting and pricing used). For the Options Menu and Random Elimination methods many of the instances had more than a single purchasing decision, as decisions data was collected iteratively as long as the subject kept purchasing the information (and there were at least two remaining values for the prize in the box). Each such iteration (corresponding to information purchasing decision with a smaller number of possible values) was taken to be an additional record in our database.

4.1 Difference in Performance

Figure 1 depicts the *IS*’s and the *IC*’s average profit over all instances experimented, according to information selling method and the price level set.² This corresponds to the scenario where the *IS* is using fixed pricing (in contrast to the dynamic pricing that will be used in following paragraphs) and serves two purposes. First, it provides evidence for a clear domination of Full Information and Options Menu over Random, as far as the seller’s profit is concerned, in all price levels. This is in contrast to the theoretical expectations, according to which all methods are alike, as proved in former sections. This can be explained by the fact that people like to be in control over the process (Tsai, Klayman, and Hastie 2008) and Random Elimination does not allow them to choose the value to be queried. Second, we can conclude from the figure that overpricing is favorable - in all three methods the average profit increases as the price level used increased, with no exceptions.³ Apparently, despite having less subjects buy the information with the higher price levels, the total profit will increase. Meaning that the increase in profit in those times where information is purchased is greater than the loss due to the decrease in the number of purchases. While obviously this will not last for substantially higher price levels, within the price level checked the $1.8 \cdot VoI$ price level is the one to be used in all three methods in order to maximize profit. Therefore if having to commit to a single information selling method and using a fixed pricing scheme, the *IS* should use the Full Information method (pricing at $1.8 \cdot VoI$), which offers the maximum profit. Interestingly, this is the combination which results in the lowest buyers’ profit. We note that for the Full Information and the Options Menu there is no statistically significant difference between the results achieved for pricing the information as $1.5 \cdot VoI$ and $1.8 \cdot VoI$. For the Random Elimination case, the difference when setting the price to $1.2 \cdot VoI$ and $1.5 \cdot VoI$ is not statistically significant. Still, the influence of

²All numerical results presented in the paper are statistically significant using t - test, $p < 0.005$, unless otherwise stated.

³The reverse relationship is observed when considering the buyer’s average profit, however this is quite intuitive as buyers always benefit from reducing the requested price.

these over the conclusions is negligible.

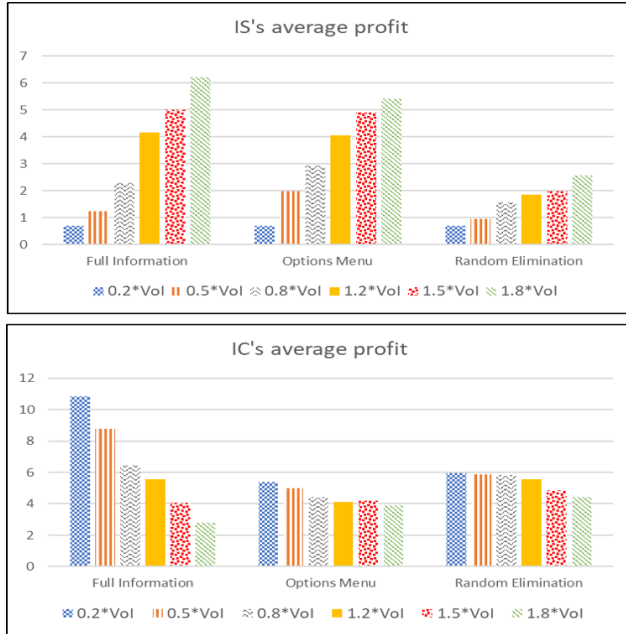


Figure 1: The *IS*'s and *IC*'s average profit when the *IS* is obligated to a fixed method and a fixed price.

4.2 Using Dynamic Pricing

Naturally, if the *IS* can use dynamic pricing, i.e., set a different price for each different problem instance, then she can substantially improve her expected profit from selling the information to the *IC*. This, however, calls for the use of machine learning methods, in order to predict the chance of purchasing the information for any price level when using each method. In this paper we use decision trees as a means for determining the price level to be used. Decision tree classification provides a rapid and useful solution for classifying instances in large datasets with a large number of variables. The use in decision trees in classification and prediction applications is very common due to its many key advantages. For example, decision trees implicitly perform variable screening or feature selection. When fitting a decision tree to a training data-set, the top few nodes on which the tree is split are essentially the most important variables within the data-set and feature selection is completed automatically. This helps a self interested *IS* to identify the set of features mostly influencing *IC*'s willingness to purchase the information, learning how to correctly predict if a purchase will be made. In addition, nonlinear relationships between parameters do not affect tree performance. Thus, we can use them in scenarios where it is known that the parameters are nonlinearly related, as in our case.

Our implementation of decision trees was done using python 3.6.4. In order to improve prediction accuracy we started with a large set of features. These includes the basic characteristics of each core problem, such as: the number of possible values, their variance, their average, the minimal

value and the maximal value. Due to the fact that people are being influenced strongly by negative values (Kahneman 2000) we also included the ratio between the number of positive and negative values for each case as a feature. For this purpose we calculated the Positive Negative Ratio (*PNR*) measurement. The calculation of the *PNR* measurement for a given case includes dividing the number of positive possible values by the number of negative possible values. One additional feature that we consider is the ratio between the value of the information and its cost. As discussed above, this ratio can be one of six possible values (0.2, 0.5, 0.8, 1.2, 1.5, 1.8). Finally, since if choosing the Random Elimination or the Options Menu options the user can choose to use them again (until left with only one value or discovering the true value), we also consider for each case, the decision number in the round. We hypothesize that for each given case, the *IC* might act differently depending on the amount of decisions she already took.

Having those features, we created a decision tree for each of the tree methods, based on the CART (Classification and Regression Trees) algorithm (Breiman 1984). For deciding how to split the data at each node we used the *Gini impurity* (Timofeev 2004). For building the tree we used *k-fold cross validation*. In *k-fold* cross validation the data is being divided into *k* equal subsets of the data called bins. Each of those bins, whose size is equal to the total data size divided by *k*, is being used as a testing set in a separate learning experiment where the tree is being built using the remaining *k - 1* bins (put together as a training set). Searching for the optimal value of *k*, i.e., the one which minimizes the error, we found that for the Full Information method, the smallest error is being received for *k = 5* whereas for the Random Elimination and Options Menu methods it is being received for *k = 10*. In order to overcome the risk for overfitting we used post-pruning (Osei-Bryson 2007). Additionally, for every learning experiment, we made sure that the users participated in the training set cases are not the same as the ones participated in the test set cases. Finally, we note, that both in the Options Menu method where the choice of which value to check is being done by the *IC*, and the Random Elimination method where a value is being randomly eliminated, there is uncertainty regarding the exact value which will be chosen. Thus, for each case, we used the average of the results achieved for all possible selections in the given case.

Figure 2 demonstrates the improvement one can achieve if using dynamic pricing using our decision tree implementation. It provides for each method the average profit of the *IS* when using the best fixed pricing level ($1.8 \cdot Vol$ according to Figure 1) compared to when having the ability to use dynamic pricing (i.e., setting the price according to core problem characteristics). From the figure we conclude that while with all three methods dynamic pricing managed to improve *IS*'s expected profit, the magnitude of improvement achieved substantially varies. With full information, the expected profit substantially improved (an increase of 52.8%) whereas with the two other methods a relatively moderate improvement was achieved. We note that the reason for the difference in the improvement's magnitude is not completely clear and requires further investigation.

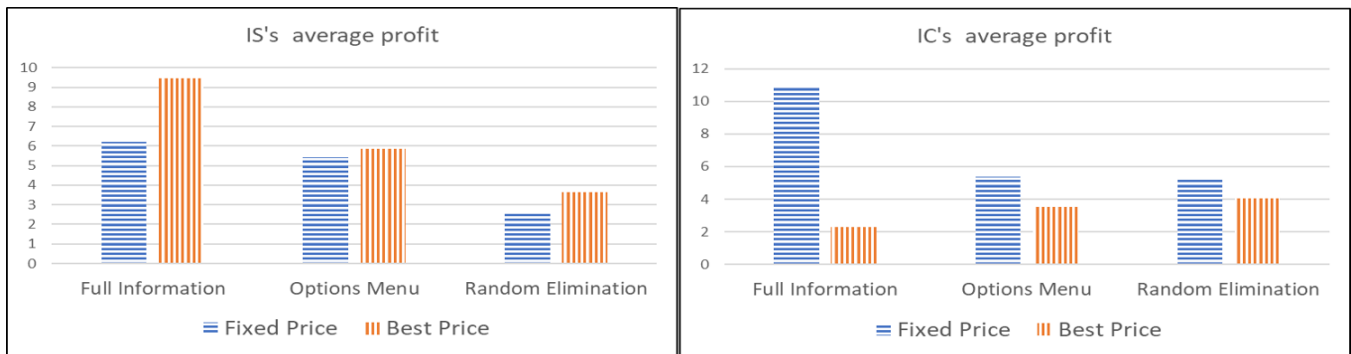


Figure 2: A comparison between the *IS* average profit and the *IC*'s average profit when the *IS* is being restricted to a fixed price and when she is using dynamic pricing.

4.3 Controlling Both Price and Method

While dynamic pricing indeed resulted in improved performance, the dominating method according to Figure 2 remains Full Information, resulting in the greatest expected profit for the *IS* by far. Still, as we show in the following paragraphs, the two other methods are important and should not be neglected. The importance of these methods arises whenever the *IS* can control both the method to be used and the price to be set for the information specifically for each core problem.

Using the above described decision trees, for each case out of the 250 core cases, the *IS* is able to compute, for each of the three methods and for each of the six prices, whether the information will be purchased or not. We emphasize that the decision trees we used for predicting the chance of purchase for any given case were only those that did not use that case in their training process, i.e., using the other $k - 1$ bins that do not include this case, as explained above. For each of the 250 core problems we picked the method and price that maximize the *IS*'s expected profit. We then checked the actual purchase decision of people in our experiments when using that method and price, replacing the average profit based on the actual decisions made over all 250 problems.

	Full Info	Random Elimination	Options Menu	Using Decision Tree
IC	2.27	4.03	3.52	2.77
IS	9.47	3.62	5.86	10.5

Table 1: A comparison of the results achieved using decision trees and not using it

Table 1 depicts the results achieved when enabling control over the and price compared to the ones achieved using each of the methods alone when the *IS* has the freedom to choose the best price out of the possible six. As observed from the table, if enabling control over both method and price, the *IS*'s expected profit is equal to 10.5, which is higher than the results achieved using each of the three different methods. Finally, we note that, as can be seen from the table, in some cases, the improvement in the *IS*'s profit comes at the expense of the *IC*. This, in many cases, might lead to some governmental prohibitions and restrictions designed to protect the *IC*. However, as can be seen from the above results,

the *IS* can easily solve this problem by paying the *IC* the difference (since even if paying the *IC* the entire loss, the *IS*'s profit is still higher if using the suggested technique).

5 Conclusions and Future Work

A pressing question when dealing with a self interested *IS*, trying to optimize her expected profit from selling the information, is how can she use wisely the information she owns in order to increase her profit from the process. In this paper we provide a new perspective to this question trying to investigate different selling methods and their combination in order to learn their influence on the seller's expected profit.

If assuming the prospective *IC*'s of the information to be rational agents, we provide a formal proof showing that the seller's expected profit from the use of each of the methods (or any combination of) is identical, i.e., the seller is indifference between all three.

When it comes to people however, this is not the case. People's bounded rationality lead to different results and therefore need to be dealt differently. Using machine learning tools, specifically decision trees, we provide a strategic *IS* with a tool she can use in order to choose a suitable sale method and price for every given scenario such that she will be able to highly improve her expected profit from the sale. This new selling technique is of much importance due to everyday situations in which people are the prospective *IC*'s of the information. In such cases, *IS*'s need to adjust their ways of actions, as shown in this paper. Finally, we note, that due to the option to compensate the *IC* for her loss, the use of the proposed method can result in practical win-win solutions.

6 Acknowledgments

This research was partially supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1162/17), the ISF-NSFC joint research program (grant No. 2240/15), the Shanghai Sailing Program (Grant No. 18YF1407900) and the National NSFC Grant 61561146398.

References

Alkoby, S., and Sarne, D. 2015. Strategic free information disclosure for a vickrey auction. In *International Workshop on Agent-*

- Mediated Electronic Commerce and Trading Agents Design and Analysis*, 1–18. Springer.
- Alkoby, S., and Sarne, D. 2017. The benefit in free information disclosure when selling information to people. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, 985–992.
- Alkoby, S.; Sarne, D.; Segal-Halevi, E.; and Sharbat, T. 2018. Eliciting truthful unverifiable information. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-18)*, 1850–1852.
- Alkoby, S.; Sarne, D.; and Das, S. 2015. Strategic free information disclosure for search-based information platforms. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-15)*, 635–643.
- Alkoby, S.; Sarne, D.; and David, E. 2014. Manipulating information providers access to information in auctions. In *Technologies and Applications of Artificial Intelligence*. Springer. 14–25.
- Alkoby, S.; Sarne, D.; and Milchtaich, I. 2017. Strategic signaling and free information disclosure in auctions. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, 319–327.
- Azaria, A.; Rabinovich, Z.; Goldman, C. V.; and Kraus, S. 2015. Strategic information disclosure to people with multiple alternatives. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(4):64.
- Azoulay-Schwartz, R., and Kraus, S. 2004. Stable repeated strategies for information exchange between two autonomous agents. *Artificial Intelligence* 154(1):43–93.
- Bergemann, D.; Bonatti, A.; and Smolin, A. 2018. The design and price of information. *American Economic Review* 108(1):1–48.
- Breiman, L. 1984. Algorithm cart. *Classification and Regression Trees*. California Wadsworth International Group, Belmont, California.
- Bro Miltersen, P., and Sheffet, O. 2012. Send mixed signals: earn more, work less. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 234–247. ACM.
- Buntain, C.; Azaria, A.; and Kraus, S. 2014. Leveraging fee-based, imperfect advisors in human-agent games of trust. In *Proceedings of the Twenty-Eight AAAI Conference on Artificial Intelligence (AAAI-14)*, 916–922.
- Cheng, H. K., and Koehler, G. J. 2003. Optimal pricing policies of web-enabled application services. *Decision Support Systems* 35(3):259–272.
- Daskalakis, C.; Papadimitriou, C.; and Tzamos, C. 2016. Does information revelation improve revenue? In *Proceedings of the 2016 ACM Conference on Economics and Computation*, 233–250. ACM.
- Dufwenberg, M., and Gneezy, U. 2002. Information disclosure in auctions: an experiment. *Journal of Economic Behavior & Organization* 48(4):431–444.
- Emek, Y.; Feldman, M.; Gamzu, I.; and Tennenholtz, M. 2011. Revenue maximization in probabilistic single-item auctions via signaling. In *Seventh Ad Auctions Workshop*. Online proceedings at <http://sites.google.com/site/adauctions2011/program-and-schedule>.
- Esó, P., and Szentes, B. 2007. Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies* 74(3):705–731.
- Ganuza, J.-J., and Penalva, J. S. 2010. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica* 78(3):1007–1030.
- Hajaj, C., and Sarne, D. 2017. Selective opportunity disclosure at the service of strategic information platforms. *Autonomous Agents and Multi-Agent Systems* 31(5):1133–1164.
- Hajaj, C.; Hazon, N.; and Sarne, D. 2015. Improving comparison shopping agents’ competence through selective price disclosure. *Electronic Commerce Research and Applications* 14(6):563–581.
- Hajaj, C.; Hazon, N.; and Sarne, D. 2017. Enhancing comparison shopping agents through ordering and gradual information disclosure. *Autonomous Agents and Multi-Agent Systems* 31(3):696–714.
- Horvitz, E.; Kadie, C.; Paek, T.; and Hovel, D. 2003. Models of attention in computing and communication: from principles to applications. *Communications of the ACM* 46(3):52–59.
- Hotz, V. J., and Xiao, M. 2013. Strategic information disclosure: The case of multiattribute products with heterogeneous consumers. *Economic Inquiry* 51(1):865–881.
- Hui, B., and Boutilier, C. 2006. Who’s asking for help?: a bayesian approach to intelligent assistance. In *Proceedings of IUI*, 186–193.
- Kahneman, D. 2000. 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.
- Lai, K. P.; Chong, S. C.; Ismail, H. B.; and Tong, D. Y. K. 2014. An explorative study of shopper-based salient e-servicescape attributes: A Means-End Chain approach. *International Journal of Information Management* 34(4):517–532.
- Moscarini, G., and Smith, L. 2003. The optimal level of experimentation. *Econometrica* 69(6):1629–1644.
- Osei-Bryson, K.-M. 2007. Post-pruning in decision tree induction using multiple performance measures. *Computers & operations research* 34(11):3331–3345.
- Paolacci, G.; Chandler, J.; and Ipeirotis, P. G. 2010. Running experiments on amazon mechanical turk.
- Rabin, M. 1998. Psychology and economics. *Journal of Economic Literature* 11–46.
- Sarne, D.; Alkoby, S.; and David, E. 2014. On the choice of obtaining and disclosing the common value in auctions. *Artificial Intelligence* 215:24–54.
- Timofeev, R. 2004. Classification and regression trees (cart) theory and applications. *Humboldt University, Berlin*.
- Tsai, C. I.; Klayman, J.; and Hastie, R. 2008. Effects of amount of information on judgment accuracy and confidence. *Organizational Behavior and Human Decision Processes* 107(2):97–105.
- Wang, Z., and Tang, P. 2015. Optimal auctions for partially rational bidders. In *Proceedings of the International Joint Conferences on Artificial Intelligence (IJCAI-15)*, 118–124.
- Yakout, M.; Elmagarmid, A. K.; Neville, J.; Ouzzani, M.; and Ilyas, I. F. 2011. Guided data repair. *Proceedings of the VLDB Endowment* 4(5):279–289.