Journal of Artificial Intelligence Research 54 (2015) 437-469 Submitted 11/14; published 11/15

Constraining Information Sharing to Improve Cooperative Information Gathering

Igor Rochlin IGOR.ROCHLIN@GMAIL.COM
School of Computer Science,
College of Management, Rishon LeZion, Israel.

David Sarne DAVID.SARNE@GMAIL.COM
Department of Computer Science,
Bar-Ilan University, Ramat-Gan, Israel.

Abstract

This paper considers the problem of cooperation between self-interested agents in acquiring better information regarding the nature of the different options and opportunities available to them. By sharing individual findings with others, the agents can potentially achieve a substantial improvement in overall and individual expected benefits. Unfortunately, it is well known that with self-interested agents equilibrium considerations often dictate solutions that are far from the fully cooperative ones, hence the agents do not manage to fully exploit the potential benefits encapsulated in such cooperation. In this paper we introduce, analyze and demonstrate the benefit of five methods aiming to improve cooperative information gathering. Common to all five that they constrain and limit the information sharing process. Nevertheless, the decrease in benefit due to the limited sharing is outweighed by the resulting substantial improvement in the equilibrium individual information gathering strategies. The equilibrium analysis given in the paper, which, in itself is an important contribution to the study of cooperation between self-interested agents, enables demonstrating that for a wide range of settings an improved individual expected benefit is achieved for all agents when applying each of the five methods.

1. Introduction

In many settings agents can benefit from cooperating in information gathering (Rochlin, Aumann, Sarne, & Golosman, 2014; Kephart & Greenwald, 2002; Rochlin, Sarne, & Mash, 2014; Hazon, Aumann, Kraus, & Sarne, 2013). For example, consider two travel agents, from the same city, that plan to participate in an international tourism conference, taking place in a highly traveled destination. There are many airlines offering flights to nearby destinations, each setting a price according to various external factors such as seat availability and agreements it has with its airlines partners. Similarly, depending on the airport of arrival, one can get to the conference by train, bus, ferry, taxi or any combination of these for different segments of the trip. Each of these means of transportation may be characterized by a different availability and fare, depending, for example, on the time of the day when it is required. Checking the feasibility and cost of the different alternatives for traveling to the conference, thus, potentially involves several time consuming activities, such as checking locations on the map and checking the companies' web-sites for routes, timetables, fares and availability, and thus incurs some "opportunity cost". Since both agents can benefit from the information each of them gathers regarding the different options for getting to the conference, they have a strong incentive to share their findings, i.e., execute the information gathering process (hereafter denoted IGP) *cooperatively*.

Cooperative information gathering is used in many real-life applications of different domains. For example, consider two friends, both interested in buying a big TV screen. The friends can visit the shopping mall, together, while each of them checks offers in different stores, and eventually they meet and share their findings. Alternatively, consider an oil drilling company sending multiple agents to explore possible drilling sites, in order to develop the best site discovered. Similarly when a position needs to be filled, HR personnel can interview candidates in parallel and recruit the best candidate found. Students can jointly look for references for an assignment they receive and eventually use the best source found by any of them.

The benefits of multi-agent cooperative information gathering are twofold. First, since each alternative (hereafter termed "opportunity") reviewed can benefit many agents, the relative cost of information gathering is reduced, while the overall welfare increases. Secondly, the task can potentially be divided according to the expertise of the different agents, if such expertise exists.[1]

Cooperative information gathering can be seen as a type of a public goods game, where all agents contribute through their individual IGP and the collective result influences the welfare of all of them. Like in public goods games, the costs of IGP in cooperative information gathering, are basically born by individual agents although benefits (better information in our case) are societal. In public goods games, in general, inefficiencies in private giving commonly occur whenever the agents are self-interested (de Jong, Tuyls, & Verbeeck, 2008; de Jong & Tuyls, 2011). Similarly, we have shown in prior work that cooperative information gathering, carried out by self-interested agents, does not result in the amount of cooperation as in the optimal fully cooperative case (Rochlin et al., 2014). In particular we have demonstrated that methods and instruments (termed "enhancers") that are easily proved to be beneficial in the fully cooperative case, can actually have a negative impact, both on individual and overall performance, in the self-interested case. These enhancers included an increase in the size of the group of agents that gather information jointly, an increase in the number of opportunities each agent has time to potentially gather information on, an improvement in some of the agents' information gathering competence, an increase in the level of heterogeneity in the individual information gathering competence of the group members and the ability to communicate throughout the process. Alas, our prior work was mostly descriptive in the sense that it outlined the potential problems that may arise in cooperative information gathering of self-interested agent. The research reported in the current paper aims to provide solutions to these problems, in the form of five somehow non-intuitive methods that essentially constrain and limit the ISP such that the individual benefit of all participating agents substantially increase.

The five cooperative information gathering methods reported in this paper differ in the constraints they put on the information sharing process (henceforth denoted ISP). The first, denoted "Enforced probabilistic information sharing" prevents individual agents from taking part in the ISP according to some probabilistic function. The second, denoted "Threshold restricted information sharing", prevents all agents that have found highly favorable values along their individual IGP from taking part in the ISP. The third, denoted "Cost filtered information sharing", introduces some cost for taking part in the ISP (where the proceeds are wasted and are not returned to the agents) and allows agents to choose whether to take part in the ISP or not. The fourth, denoted "Random finding sharing", allows all agents to take part in the ISP, however restricts each of them to disclosing only one of the values from the set known to each, in random. Finally, the fifth, denoted "Subgroup restricted information sharing", initially divides the agents into subgroups and allows only local ISPs,

1. For buyers' cooperation, the agents can also benefit from a volume discount through their cooperation; however this property holds only for that specific domain.

i.e., information sharing in the subgroup level. Each of these methods may seem counter intuitive, because the absence of some of the agents in the ISP or the restriction on the amount of information that can be shared could be harmful to all agents. Yet, in many settings, the use of these methods can be highly beneficial. This is because of the paradox embedded in the ISP option - while the sharing of information benefits all agents, the very fact that all information gathered is going to be shared, discourages agents from investing much resources in their individual IGP (Rochlin et al., 2014). Therefore, with the use of these methods the individual benefit of each agent from taking part in the ISP decreases, however the IGP carried out by the agent individually becomes more efficient. Therefore, by intelligently managing the tradeoff between the two, a more beneficial equilibrium can be achieved, which improves both the overall and individual benefits.

The paper provides a comprehensive analysis of the individual information gathering strategies used by the agents, given the strategy of others, under the different methods. For the Enforced probabilistic information sharing, Random finding sharing and Subgroup information sharing methods the agents' individual strategy is proven to be similar in structure to the one used with the standard cooperative information gathering method — the agent will resume information gathering as long as the best value obtained so far is lower than some reservation value (a threshold), regardless of how much more information can potentially be gathered. For the Threshold restricted information sharing method the agents' individual strategy is proven to be threshold-based, however the threshold changes as a function of the amount of information that potentially can still be gathered. For the Cost filtered information sharing method, the individual strategies are proven to be based on a single reservation value for determining the benefit in additional information gathering and a set of intervals for deciding whether to take part in the ISP. These allow the characterization of the resulting equilibria. Using synthetic environments, we numerically demonstrate that all five methods result in substantial improvement to each of the agents' individual expected benefit for a wide range of settings.

The results contribute to the advancement of theories of cooperation in MAS. As discussed later in the paper, the methods can be easily applied and their use can benefit both individuals planning to engage in cooperative information gathering and designers of multi-agent systems (MAS) where cooperative information gathering is likely to take place.

In the following section we formally introduce the cooperative information gathering model. In Section 3 we detail the model analysis, the equilibrium strategies for the different model variants considered and supply numerical examples for the benefit that can be achieved when using them. Related work is reviewed in Section 4, emphasizing the uniqueness of the analysis provided in the paper. Finally, we conclude and discuss directions for future research in Section 5.

2. The Model

The model considers a set $K = \{A_1, ..., A_k\}$ of fully-rational self-interested agents. Each of the agents needs to gather information pertaining to the value (e.g., benefit) of different opportunities to which it has access and eventually choose one. The values of the different opportunities are a priori unknown and information is gathered for one opportunity at a time. The individual information gathering problem, as defined above, is standard and follows the assumptions commonly used in literature (Chhabra & Das, 2011; Kephart & Greenwald, 2002; Hazon et al., 2013; Rothschild, 1974; McMillan & Rothschild, 1994; Morgan & Manning, 1985). The uncertainty associated with the value of opportunities available to any agent A_i is modeled, as in most costly information gathering literature (McMillan & Rothschild, 1994; Burdett & Malueg, 1981; Carlson & McAfee, 1984; Lippman & McCall, 1976; Morgan, 1983), through a probability distribution function (p.d.f.) $f_i(x)$ (i.e., the value of each opportunity in the individual IGP of A_i is drawn from $f_i(x)$), with which all agents are familiar (Tang, Smith, & Montgomery, 2010; Waldeck, 2008; Janssen, Moraga-Gonzalez, & Wildenbeest, 2005). Due to the resource consuming nature of the process it is considered costly in the sense that revealing the value of an opportunity incurs a fixed cost, denoted c_i. The model assumes that any agent A_i is constrained by a number of opportunities accessible to this agent, denoted n_i. The cost c_i, the distribution $f_i(x)$ and the number of opportunities n_i, are defined in the agent's level to support settings where different agents have different skills and capabilities. The agent thus needs to gather information, i.e., explore the value of some of the opportunities and eventually pick one of the values revealed (i.e., recall is permitted) (Carlson & McAfee, 1984; McMillan & Rothschild, 1994).

In settings where all opportunities are applicable to all agents the agents have an incentive to cooperate in information gathering in the sense that all individual findings are eventually shared with all others. While there are many ways to share the information, the focus of this paper is on setups where the ISP takes place at some pre-specified time, after all agents have completed their individual IGPs and each needs to decide on the opportunity it will choose. The choice of sharing findings at the end of the individual IGPs is mostly natural and customary in real life. More importantly, the alternative of sharing information throughout the process has a major setback in the sense that each individual agent finds information sharing to be beneficial only when it is on the receiving end, i.e., it is the one being informed that a "favorable" opportunity has been found; when it is on the reporting end, the agent loses from such communication since the report can potentially encourage the other agents to terminate their individual information gathering. On the other hand sharing the information after concluding the individual IGPs is always beneficial for the agent as it gains more information, and at the same time the information it discloses does not affect the behavior of others thereafter since they also have already concluded their IGPs.

As in prior models of cooperative information gathering, we also assume that: (a) the agents are truthful in the sense that they always report the true values they obtain;[2] (b) each agent A_i has some fall-back value v_0^i, i.e., even if not becoming acquainted with any opportunity values (in case of not gathering any information individually and not receiving any of the others' findings) the agent can presumably benefit v_0^i;[3] and (c) either the opportunities each agent can check are unique or the agents can a priori divide the opportunities among them such that each will be assigned a different set. It is assumed that information gathering costs and opportunity values are additive and each agent A_i is interested in maximizing its expected benefit, denoted EB_i. The benefit of an agent is therefore the best value obtained by the group minus the costs accumulated individually along the agent's individual IGP. Finally, it is assumed that when engaging in a cooperative information gathering process all agents are a priori acquainted with the probability distribution functions and information gathering costs of all agents, i.e., the only possible difference in the information available to the different agents throughout the cooperative IGP is their own findings and the findings of others.

The cooperative information gathering model as detailed above can be found in full or with some variations in prior literature (Hazon et al., 2013; Rochlin et al., 2014; Gatti, 1999; Carlson &

2. The truthfulness assumption is commonly justified by a substantial potential reputation loss, and is easily enforceable using fines.
3. And similarly, if taking part in the ISP then necessarily disclosing that value in the absence of any better one.

Application	Individual goal	Opportunity	Value	Information gathering cost	Source of uncertainty
Product acquisition	Minimize individual expense	Complex service/product	Cost of purchase	Time spent finding options and evaluating them	Sellers' competition, seasonal effects, service constraints
Choosing an oil-drilling site	Maximize oil revenues minus cost of exploratory drills	Potential drilling sites	Amount of oil found	Time and resources spent in the exploratory drills	Uncertainty regarding the amount of oil in each drilling site
R&D	Minimize the cost of production and R&D expenses	Production technology	Cost of production with a specific technology	R&D cost of specific technology	Uncertainty concerning implementation aspects of a desired technology
Information Search (student's assignment)	Maximize individual utility (grade minus individual effort)	Information source (e.g., online, textbook, library resource)	Expected grade if this source is used	Time spent evaluating different sources	Differences in coverage of the topic, relevance, accuracy, level of details

Table 1: The mapping of different applications to the cooperative information gathering problem.

McAfee, 1984).[4] Taking the travel agents example, given in the previous section, the opportunities represent different alternatives for reaching the conference location and their value is their total cost. The information gathering cost is the agents' cost of the time needed to explore the alternatives. The goal of each agent is to minimize her expected expense, defined as the cost of the best alternative found by the two plus the cost of the time spent individually to review the different alternatives. Similarly, the model can be mapped to all the other applications mentioned in the introduction (e.g., see Table 1).

3. Analysis

We divide the analysis according to the five cooperative information gathering enhancing methods. For each method, first we determine the individual optimal information gathering strategy of an agent taking part in the process, as the best response to the other agents' strategies. Then, we show how the collective behavior is derived and extract the equilibrium set of strategies.

The appropriate equilibrium concept depends on how we define the type space. If we were to define an agent's type by the specific vector of values it would encounter if fully exhausting its IGP, then the appropriate concept would be the ex-ante Bayesian Nash equilibrium, since agents are a priori unaware of their types and this information is revealed to them (individually) along the IGP. However, we prove in this paper that in this definition, an agent's type would not affect its strategy (strategies turn out to be based on thresholds that are set prior to conducting the IGP). Therefore, while outcomes are stochastic, one could in theory build a direct stochastic mapping from individual strategies to the global outcome and therefore there is no need to go through the type in doing so. Therefore there is no added value here from the Bayesian Nash equilibrium characterization. The equilibrium concept would then be subgame-perfection or Stackelberg (where the system designer sets the rules for participation in the ISP, and then agents respond by choosing their optimal strategies). All these result in a similar equilibrium characterization in our case.

4. While some model variants consider the task to be executed by a representative agent, acting on behalf of the group, the essence of gathering costly information and trading-off costs and benefit is the same.

Finally, we demonstrate how the expected individual benefit of all agents improves, when the method is used, compared to the standard cooperative IGP. In order to illustrate the performance achieved with the different methods, we use a setting where the agents are homogeneous in terms of their information gathering environment. Meaning that each agent samples opportunities from the same probability distribution function $f(y)$ (i.e., $f_1(y) = .. = f_k(y) = f(y)$), that all agents are constrained to the same number of opportunities n they can sample overall (i.e., $n_1 = .. = n_k = n$) and they all share the same information gathering cost c (i.e., $c_1 = .. = c_k = c$) and same fallback v_0 (i.e., $v_0^1 = .. = v_0^k = v_0$). Such setting is quite common in real-life as often (and especially in the Internet age) people can potentially access the same opportunities and with a similar effort (e.g., time spent navigating through a web-site). For example, in our travel agents running example, it is likely that all agents have access to the same web-sites and resources that need to be used for identifying and gathering information on the different options available for getting to the conference, and that none of them, being a trained and experienced travel agent, has some specific advantage in doing so. For simplicity and ease of exposition of the figures, we use $f(y)$ to be the uniform distribution function (between 0 and 1). We stress that even though such a homogeneous setting is standard in costly information gathering literature (McMillan & Rothschild, 1994; Lippman & McCall, 1976), and common in real-life ISP as argued above, its use in our case is merely for illustration purposes and all the results concerning individual strategies and equilibrium structures that are given in this paper are based on formal theoretical proofs.

3.1 Enforced Probabilistic Information Sharing

In this method each agent A_i is a priori assigned some probability P_i^{IS} which, after it has completed its individual IGP, is used to determine whether it is allowed to take part in the ISP. It is assumed that all agents are a priori acquainted with the probabilities P_i^{IS} used for enabling information sharing of each agent. The determination whether an agent will take part in the ISP must be made in proximity to the time information should actually be shared and requires some enforcing mechanism since once the individual IGP is completed, agents obviously will benefit from taking part in the ISP, as it does not incur any cost and at the same time can improve their best finding. Furthermore, since the process takes place after all individual IGPs have been terminated, the information an agent discloses has no influence over the individual IGP strategies used by the other agents. Such enforcement is easy to achieve through simple means. For example, in our travel agents running example this is equivalent to having all agents send their findings to a designated secured server. The server will select those eligible for information sharing, according to the pre-defined probabilities, and will remove from its database all the information coming from those that are not. Then, the server will allow those eligible for information sharing access to the data it stores.

An agent's state throughout its individual IGP is represented by the subset of opportunities on which it has already gathered information, and their associated values, and consequently the remaining opportunities for which the values are still unknown. An agent's strategy is thus the mapping from a world state to a choice $\{resume, terminate\}$ where $resume$ suggests that the agent needs to gather information about an additional opportunity (a random one, since all opportunities available to a given agent are a priori alike) and $terminate$ means that the agent needs to proceed to the ISP. Theorem 1 proves that the state representation in this case can be compacted to the best value found so far (including the fallback v_0^i), v, and that the optimal strategy can be represented in terms of a single reservation value, independent of the number of remaining opportunities.

Theorem 1. *Given the probability distribution function of the maximal value obtained by all other agents that take part in the ISP, denoted $\bar{f}_i(x)$, agent A_i's optimal individual information gathering strategy is to set a reservation value $r_i \geq v_0^i$, where r_i is the solution to:*[5]

$$c_i = P_i^{IS} \cdot \int_{y=r_i}^{\infty} f_i(y) \int_{x=-\infty}^{\infty} (\max(y,x) - \max(r_i,x))\bar{f}_i(x)dxdy \tag{1}$$
$$+ (1 - P_i^{IS}) \cdot \int_{y=r_i}^{\infty} (y - r_i) f_i(y)dy$$

The agent should always choose to gather information on an additional opportunity (if one is available) if the best value obtained so far is below r_i and otherwise it should proceed to ISP.

Proof. See appendix A.

Theorem 1 specifies the optimal strategy of an agent given the strategy of others. The solution of a set of k equations similar to (1), one for each agent A_i, will provide a set of pure equilibria of the form $\{r_i | 1 \leq i \leq k\}$ if any exist. A mixed equilibrium in our case is defined by a probability $p_i(v,j)$ assigned to each state (v,j), defining whether the agent will resume or terminate information gathering in that state. This may seem infeasible to extract, based on the infinite number of states (as value distributions are continuous). Nevertheless, in order for such a solution to hold, the agent's expected benefit from both actions (resume and terminate information gathering when in that state) must be equal. Based on the optimality of the reservation-value rule, this can hold only for states where the value v equals r_i as calculated according to (1). However, due to the continuous nature of v, the probability of actually reaching states that satisfy the above condition is zero, thus assigning such probabilities will have no effect on the other agents. The only exception for the above is the agent's strategy at the beginning of its individual IGP. Here, the state is a priori known to be $(v_0^i, 0)$ hence adding some probability for actually gathering information on one opportunity and then continuing according to r_i (or otherwise going straight to ISP due to the indifference to resuming or terminating information gathering) will have an actual effect on the others. Consequently, a mixed equilibrium for our problem is of the form:

$$\{(p_i, r_i) | 1 \leq i \leq k\}$$

where p_i is the probability that agent A_i will initiate its individual IGP ($0 \leq p_i \leq 1$) and r_i is the reservation value to be used by the agent. A solution will be considered stable (i.e., in equilibrium), if none of the agents will find it beneficial to deviate from it individually. An equilibrium in pure strategies for the problem would require $p_i \in \{0,1\}$ $\forall i$. Any other solution is a mixed equilibrium strategy.

Now that the individual strategy in equilibrium has been defined in its complete form (i.e., including the probabilistic aspect), we can formulate $\bar{f}_i(x)$ (the probability distribution function of the maximal value obtained along the IGP of all other agents that will take part in the ISP). For this purpose we make use of the probability that the maximum value that will be found by all the agents that take part in the ISP, except A_i, will be smaller than or equal to x, denoted $\bar{F}_i(x)$. The

5. Notice that if the equation has a solution then it is necessarily of the form $v_0^i \leq r_i$. In case there is no solution to the equation (e.g., in case of a substantial c_i), onwards, then we can simply set $r_i = v_0$ and no information gathering will take place.

calculation of $\bar{F}_i(x)$ makes use of the probability that the maximum value obtained along the IGP of an agent A_j (that chooses to engage in IGP and uses r_j), is less than x, denoted $F_j^{\text{return}}(x)$, calculated according to:[6]

$$F_j^{\text{return}}(x) = \begin{cases} 0 & x < v_0^j \\ F_j(x)^{n_j} & v_0^j \leq x \leq r_j \\ F_j(r_j)^{n_j} + \frac{1-F_j(r_j)^{n_j}}{1-F_j(r_j)}(F_j(x) - F_j(r_j)) & x > r_j \end{cases} \tag{2}$$

The case where $x < v_0^j$ is trivial, as v_0^j is a lower bound for the best value the agent ends up with. For the case where $v_0^j \leq x \leq r_j$, the value of all n_j opportunities must result in a value below x. When $x > r_j$ there are two possible scenarios. The first is where all n_j opportunities result in a value below the reservation value r_j, i.e., with a $F_j(r_j)^{n_j}$ probability. The second, is where the information gathering terminates right after revealing value y at the lth opportunity such that $r_j < y < x$ (otherwise, if $y < r_j$ the information gathering should resume) and all the former $l-1$ values obtained are smaller than r_j (otherwise the lth opportunity is not reached). The probability of the latter case occurring (summing over all values of $l \leq n_j$) can be calculated using the geometric series $\sum_{l=1}^{n_j}(F_j(x) - F_j(r_j))F_j(r_j)^{l-1} = \frac{1-F_j(r_j)^{n_j}}{1-F_j(r_j)}(F_j(x) - F_j(r_j))$.

The probability distribution function of the maximum value obtained throughout agent A_i's IGP, denoted $f_i^{\text{return}}(x)$, is by definition, the first derivative of $F_i^{\text{return}}(x)$:

$$f_i^{\text{return}}(x) = \frac{d(F_i^{\text{return}}(x))}{dx}$$

Thus, we can now formulate the probability that the maximum value that will be found by all the agents taking part in ISP, except A_i, will be smaller than or equal to x, $\bar{F}_i(x)$:

$$\bar{F}_i(x) = \prod_{A_j \in K \wedge j \neq i} (P_j^{IS}(p_j F_j^{\text{return}}(x) + (1-p_j)) + (1 - P_j^{IS})) \tag{3}$$

The calculation is based on the probability that any given agent will either: (a) not take part in the ISP (hence it will not contribute any value) or, (b) take part in the ISP and its best value obtained in its individual IGP will be below x. The probability of (a) is $(1 - P_i^{IS})$. To calculate the probability of (b) we first need to calculate the probability that an agent's best value will be below x. This can happen either if the agent initiated IGP and the best value it will obtain will be lower than (or equal to) x, i.e., with a probability of $p_j F_j^{\text{return}}(x)$, or if the agent opted out from IGP, i.e., with a probability of $1 - p_j$. Therefore the probability of b is given by $P_j^{IS}(p_j F_j^{\text{return}}(x) + (1-p_j))$. Consequently, the probability distribution function $\bar{f}_i(x)$ is the derivative of $\bar{F}_i(x)$:

$$\bar{f}_i(x) = \frac{d\bar{F}_i(x)}{dx}$$

These enable us to calculate the expected benefit of agent A_i when the other agents use the set of strategies $\{(p_i, r_i) \,|1 \leq i \leq k\}$. If agent A_i chooses to engage in IGP then its expected benefit,

6. The maximum value found includes also the fallback v_0^j. In the degenerate case where $F_j(v_0^i) = 0$ we use $F_j^{\text{return}}(x) = 0$ for $v_0^i \leq x \leq r_j$.

denoted $EB_i(IGP)$, is given by:[7]

$$EB_i(IGP) = P_i^{IS} \cdot \int_{y=-\infty}^{\infty} f_i^{\text{return}}(y) \int_{x=-\infty}^{\infty} \max(v_0^i, y, x)\bar{f}_i(x)dxdy + \quad (4)$$
$$(1 - P_i^{IS}) \cdot \int_{y=-\infty}^{\infty} \max(v_0^i, y) f_i^{\text{return}}(y)dy - c_i \frac{1 - F_i(r_i)^{n_i}}{1 - F_i(r_i)}$$

where the first term on the right hand side is the expected maximum between the best value found by the agent itself (i.e., associated with a distribution $f_i^{\text{return}}(y)$) and the best value returned by the other agents (associated with a distribution $\bar{f}_i(x)$) if agent A_i participates in the ISP (i.e., with a P_i^{IS} probability). The second term is the expected "best" (i.e., maximum) opportunity-value found by the agent along its information gathering if the agent is not allowed to take part in the ISP (i.e., with a $1 - P_i^{IS}$ probability). The last term is the expected cost incurred throughout the IGP carried out by A_i, calculated as: $c_i \sum_{j=1}^{n_i} (F_i(r_i))^{j-1} = c_i \frac{1-F_i(r_i)^{n_i}}{1-F_i(r_i)}$, as the number of opportunities on which information is gathered is a geometric random variable bounded by n_i, with a $1 - F_i(r_i)$ success probability — the IGP terminates only upon receiving a value greater than r_i (or all n_i opportunities explored) and the probability for a value greater than r_i is $1 - F_i(r_i)$.

When the agent opts not to execute individual IGP at all, its expected benefit, denoted $EB_i(\neg IGP)$, is simply the expected value of the maximum value returned by the other agents, taking part in the ISP, if it takes part by itself in the process, or otherwise v_0^i, i.e.:

$$EB_i(\neg IGP) = P_i^{IS} \cdot \int_{x=-\infty}^{\infty} \max(v_0^i, x)\bar{f}_i(x)dx + (1 - P_i^{IS}) \cdot v_0^i \quad (5)$$

At this point, we have everything that is needed to formulate the equilibrium stability conditions. A set of strategies $\{(p_i, r_i) | 1 \leq i \leq k\}$ will be in equilibrium only if the following conditions hold:

(a) For every agent A_i for which $p_i = 0$, $EB_i(IGP) \leq EB_i(\neg IGP)$.

(b) For every agent A_i for which $p_i = 1$, $EB_i(IGP) \geq EB_i(\neg IGP)$.

(c) For every agent A_i for which $0 < p_i < 1$, $EB_i(IGP) = EB_i(\neg IGP)$.

Therefore, in order to find the equilibrium, the stability of 3^k possible solutions of type $\{(p_i, r_i) | 1 \leq i \leq k\}$ differing in the value each p_i obtains ($p_i = 0$, $p_i = 1$ and $0 < p_i < 1$) needs to be checked. For every combination, the reservation values of the different agents and the probability p_i of each agent that uses a non-pure mixed strategy (i.e., with $0 < p_i < 1$) should be calculated by solving a set of equations of type (1) (one for each agent characterized by $p_i = 0$) and $EB_i(IGP) = EB_i(\neg IGP)$ (one for every agent A_i for which $0 < p_i < 1$). Once the appropriate reservation values and probabilities are obtained for a given set, the stability conditions need to be validated.

We note that there is no guarantee that an equilibrium will actually exist (either pure or mixed, since there are an infinite number of strategies). Also, there is no guarantee that if one exists there will be no other equilibria (i.e., multiple equilibria may exist). In the latter case, if there is one

7. If none of the other agents engage in IGP (i.e., $p_j = 0$ for all $A_j \neq A_i$) then $\bar{f}_i(x) = 0$ and therefore the expected benefit $EB_i(IGP)$ should be calculated as $EB_i(IGP) = \int_{y=-\infty}^{\infty} \max(v_0^i, y) f_i^{\text{return}}(y)dy - c_i \frac{1-F_i(r_i)^{n_i}}{1-F_i(r_i)}$, where r_i is the solution to $c_i = \int_{y=r_i}^{\infty} (y - r_i) f_i(y)dy$ (according to a single agent's optimal IGP (Rochlin et al., 2014)), and $EB_i(\neg IGP) = v_0^i$.

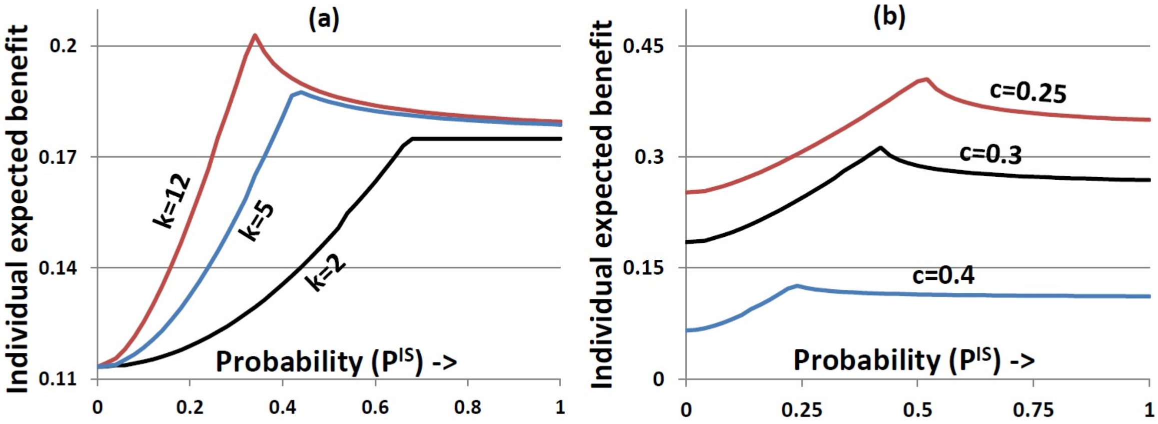

Figure 1: Enforced probabilistic information sharing - effect of P^{IS} on the individual expected benefit, for different: (a) numbers of agents, k, in a setting: $c = 0.35$ and $n = 5$; and (b) information gathering costs, c in a setting: $k = 15$ and $n = 4$.

equilibrium that dominates the others in terms of the individual expected benefit each and every agent obtains then it will likely be the one used. Otherwise, there is no way of deciding which of the equilibria is the one to use, and this question is not included in the scope of the current paper.

We emphasize that the above analysis generalizes the analysis of the standard cooperative information gathering model (Hazon et al., 2013; Rochlin et al., 2014) in the sense that the latter is a specific case of the first, where the probability each agent will be allowed to take part in the ISP is one (i.e., $P_i^{IS} = 1 \,\forall 1 \leq i \leq k$). Furthermore, when the probability each agent will be allowed to take part in the ISP is zero, the solution obtained is the same as the one known for the single-agent information gathering problem (McMillan & Rothschild, 1994) (since each agent relies solely on the values it obtains throughout its individual IGP).

Figure 1 depicts the agents' individual expected benefit as a function of the probability P_i^{IS} used, for different group sizes (k) and information gathering costs (c). The setting used is the homogeneous setting described at the beginning of the section and the value of P_i^{IS} is the same for all agents (i.e., $P_i^{IS} = P^{IS} \,\forall i$). The other model parameters were set to: $c = 0.35$ and $n = 5$ (Figure 1(a)) and $k = 15$ and $n = 4$ (Figure 1(b)). As depicted in the figure, the maximum expected benefit (agent-wise, as all agents are alike in this case) is obtained when the participation of the agents in the ISP is not certain but rather determined probabilistically (i.e., $0 < P^{IS} < 1$). The typical pattern exhibited in the figure is an increase and then a decrease in the expected individual benefit as the probability P^{IS} increases. This is explained as follows. When $P^{IS} = 0$ each agent actually executes an individual IGP without any information sharing with others. As P^{IS} increases, the agent relies more on other agents' findings. Thus, p_i and r_i become lower, which is bad for the group since everybody gains less from the participation of the agent in the ISP. However, at the same time the probability the agent will actually take part in the ISP increases, thus, overall, the individual expected benefit increases. Nevertheless, for some value of P^{IS}, the loss due to the resulting decrease in p_i and r_i becomes more dominant than the benefit due to the increase in the value of P^{IS}.

Another interesting behavior observed in Figure 1 is that the increase in information gathering costs, c, and the increase in the number of agents, k, results in a decrease in the the value of P^{IS} that maximizes expected benefit. This may seem non-intuitive since the greater the information gathering cost the greater the potential benefit that can be achieved by information sharing and similarly, the greater the number of agents the greater the chances of obtaining favorable values in the ISP. Therefore further limiting information sharing in settings with high k and c values may seem unnatural. The phenomena is explained by the fact that the positive effect of the increase in P^{IS} over the participation probabilities p_i and the reservation value r_i used by each agent in equilibrium in this case, which are substantially poor to begin with, is greater than the loss due to the uncertain information sharing.

While the above method requires that an agent either completely avoid or fully participate in the ISP, many other variants can be considered, e.g., partially limiting information sharing. For example, an agent can be requested, with some probability, to contribute the information it has gathered, without receiving the other group members' information. Alternatively, an agent can be allowed to receive information however not share its own gathered data. These variants can, in some settings, result in substantially superior performance as we illustrate in Appendix B.

3.2 Threshold Restricted Information Sharing

With this method the agents that have found highly favorable values along their individual IGP are prevented from taking part in the ISP. This is implemented by setting a threshold V_i^{IS}, for each agent A_i, which requires that the agent opt-out of the ISP if it obtained a value greater than V_i^{IS} in its individual IGP. It is assumed that all agents are a priori acquainted with the thresholds of all agents. In our travel agents running example this can take the same form as with the Enforced probabilistic information sharing, with slight modifications, e.g., having all agents send their findings to a designated secured server. The server will select those eligible for information sharing, according to the pre-set thresholds, and will remove from its database all the information coming from those that are not. Then, the server will allow those eligible for information sharing access to the data it stores.

The choice of excluding agents with findings that are greater than the threshold may seem counter intuitive. Seemingly those with favorable findings, i.e., those that actually performed well and can contribute most to the others are being punished. One may argue that a more suitable choice would be to set the threshold such that agents that have not found good values would be excluded from the ISP, thereby encouraging them to try harder. Nonetheless, this also suggests some discouragement to agents in their IGPs as they now will have greater expectations that high values will be reported in the ISP. Obviously, at the time the ISP ought to take place, opting out is dominated by taking part in the ISP, both individually and globally. All agents will benefit from the participation of more agents in the ISP. Moreover, the fact that we require that agents with more favorable values not to take part in the ISP seems to intensify the potential negative effect of this method as discussed above. Nevertheless, since each of the agents may find itself in a situation where it is requested not to take part in the information sharing, and due to the decrease in the potential for improvement encapsulated in information sharing to those agents that will eventually take part in it, it is likely that each agent will be individually motivated towards a more efficient IGP.

Overall, the Threshold restricted information sharing method is a bit more complicated to enforce since agents may choose to declare a value different than the best they have found, in an effort

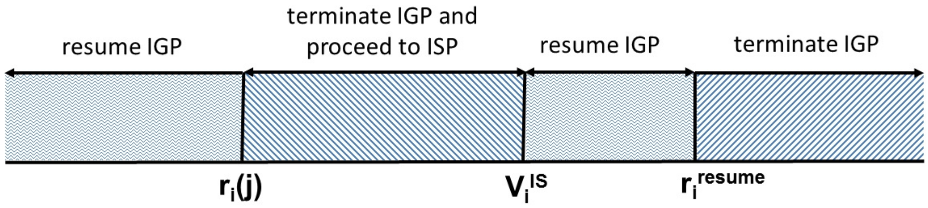

Figure 2: A schematic illustration of the optimal strategy for Threshold restricted information sharing.

to take part in information sharing. Nevertheless, whenever the true value can be validated (e.g., in the travel agents example, each agent can be requested to post a receipt indicating the way she got to the conference and the amount paid, after the ISP) then the method can be easily enforced using fines.

In this case, as we prove in the following paragraphs, the agent's strategy needs to take into consideration not only the best value obtained thus far, but also the number of opportunities on which information has already been gathered. This is because whenever a value $v > V_i^{IS}$ is encountered, whereby the agent is excluded from the ISP, the agent can still benefit from resuming its individual IGP, and the benefit from additional information gathering depends on the number of remaining opportunities.

The structure of each agent's optimal strategy, given the strategy of others, is given in Theorem 2.

Theorem 2. *Given the probability distribution function of the maximal value obtained by all other agents that take part in the ISP, $\bar{f}_i(x)$, and the number of opportunities on which information has already been gathered, $j \leq n_i$, agent A_i's optimal individual information gathering strategy can be described by the pair $(r_i(j), r_i^{resume})$, such that $v_0^i \leq r_i(j) \leq V_i^{IS} \leq r_i^{resume}$ (see Figure 2), where:*

$$r_i^{resume} = \max\left(V_i^{IS}, r | c_i = \int_{y=r}^{\infty} (y-r) f_i(y) dy\right) \tag{6}$$

and $r_i(j)$ is the solution to:

$$\begin{aligned} c_i = & \int_{y=r_i(j)}^{V_i^{IS}} f_i(y) \int_{x=-\infty}^{\infty} (\max(y,x) - \max(r_i(j),x)) \bar{f}_i(x) dx dy \\ & + \int_{y=V_i^{IS}}^{\infty} f_i(y) \int_{x=-\infty}^{\infty} (EB_i^{j+1}(y) - \max(r_i(j),x)) \bar{f}_i(x) dx dy \end{aligned} \tag{7}$$

where the $EB_i^j(y)$ is given by:

$$EB_i^j(v) = \begin{cases} v & r_i^{resume} \leq v \vee j > n_i \\ -c_i + \int_{y=-\infty}^{\infty} EB_i^{j+1}(\max(v_0^i, y, v)) f_i(y) dy & v < r_i^{resume} \wedge j \leq n_i \end{cases} \tag{8}$$

Given the best value obtained so far, v, Agent A_i should resume its IGP if $v < r_i(j)$ or $V_i^{IS} < v \leq r_i^{resume}$ and otherwise terminate (and proceed to ISP if $v \leq V_i^{IS}$).

Proof. See appendix C.

Intuitively, the reservation value r_i^{resume} is used to determine if the IGP should be resumed in cases in which a favorable value $v > V_i^{IS}$ has been found and the agent will now be "on its own". The reservation value $r_i(j)$ is used to determine if the IGP should be resumed when $v \leq V_i^{IS}$, i.e., when the agent can still potentially take part in the ISP. Resuming the IGP in the latter case can lead to better values, however at the same time can also lead to exclusion from the ISP, in which case the agent will end up on its own. Since the fallback in case of a finding $v > V_i^{IS}$ depends on the number of remaining opportunities j, we use a different reservation value $r_i(j)$ for each j value. In case $r_i^{resume} = V_i^{IS}$, there is no benefit for the agent to resume the IGP once a value $v > V_i^{IS}$ is found. In this case $r_i(j) = r$ for any j.

Once again, the solution of a set of equations consisting of (6-8) will provide a set of pure equilibria of the form $\{(r_i(1), .., r_i(n_i)), r_i^{resume} | 1 \leq i \leq k\}$, if any exist. Similarly, for the same considerations given in Section 3.1, a mixed equilibrium for this case will be of the form: $\{(p_i, (r_i(1), .., r_i(n_i)), r_i^{resume}) | 1 \leq i \leq k\}$.

While the equilibrium analysis in this case generally resembles the one given in 3.1, the calculation of $F_i^{\text{return}}(x)$, which is the probability that the maximum value obtained by agent A_i's individual IGP (including v_0^j) will be less than x is substantially more complicated. In order to formulate $F_i^{\text{return}}(x)$ we use $F_i^{\text{return}}(x, v, j)$ to denote the probability that agent A_i will obtain a maximum value of x or below, given its current state (v, j), where v is the best value obtained so far in its IGP (including v_0^i) and j is the number of opportunities on which information has been gathered. The function $F_i^{\text{return}}(x, v, j)$ can be calculated recursively according to:

$$F_i^{\text{return}}(x, v, j) = \begin{cases} 0 & v > x \\ 1 & v \leq x \wedge \left(r_i(j) < v \leq V_i^{IS} \vee r_i^{resume} \leq v \vee j = n_i\right) \\ \int_{y=-\infty}^{\infty} F_i^{\text{return}}(x, \max(v_0^i, y, v), j+1) f_i(y) dy & \text{otherwise} \end{cases} \tag{9}$$

The case where $x < v$ is trivial since the maximum value the agent will obtain will be at least v. Therefore, the probability of obtaining x or below is zero. Similarly, when $r_i(j) < v \leq V_i^{IS}$ or $r_i^{resume} \leq v$ or when there are no additional opportunities ($j = n_i$) the agent will inevitably terminate its IGP (according to Theorem 2) and the maximum value that will be obtained will be v. In such case, if $v \leq x$ the function obtains 1. In all other cases, the IGP resumes, hence the probability is given recursively based on the new state $(x, \max(v, y), j + 1)$ the agent will be in after gathering information on one additional opportunity. Using $F_i^{\text{return}}(x, v, j)$ we can calculate the probability that the maximum value that will be obtained by agent A_i's individual IGP will be less than (or equal to) x, as: $F_i^{\text{return}}(x) = F_i^{\text{return}}(x, -\infty, 0)$.

Thus, we can now formulate the probability that the maximum value provided by agent A_i in the ISP will be less than x, denoted $F_i^{\text{return}'}(x)$:

$$F_i^{\text{return}'}(x) = \begin{cases} 0 & x < v_0^i \\ F_i^{\text{return}}(x) + (1 - F_i^{\text{return}}(V_i^{IS})) & v_0^i \leq x \leq V_i^{IS} \\ 1 & \text{otherwise} \end{cases} \tag{10}$$

The case where $x < v_0^i$ is trivial, as v_0^i is a lower bound for the best value the agent ends up with. The case where $x \leq V_i^{IS}$ is satisfied either if A_i's best value is less than x (in which case it will take part in the ISP since $x \leq V_i^{IS}$), i.e., with a probability of $F_i^{\text{return}}(x)$, or if the agent will not take part in the ISP, i.e., with a probability of $1-F_i^{\text{return}}(V_i^{IS})$. The case where $x > V_i^{IS}$ is straightforward — since agent A_i will take part in the ISP only if its best value will be lower than V_i^{IS}, the probability of obtaining a value v (where $x > V_i^{IS}$) from it is 1.

Using $F_i^{\text{return}'}(x)$, we can calculate the function $\bar{F}_i(x)$ (a modification of (3)):

$$\bar{F}_i(x) = \prod_{A_j \in K \wedge j \neq i} (p_j F_j^{\text{return}'}(x) + (1 - p_j)) \tag{11}$$

The probability distribution function $\bar{f}_i(x)$ is the first derivative of $\bar{F}_i(x)$ as before. Similarly, the probability distribution function $f_i^{\text{return}}(x)$ is the first derivative of $F_i^{\text{return}}(x)$.

We now turn to calculating the expected cost incurred throughout the IGP carried out by agent A_i, denoted EC_i. In order to calculate EC_i we use $EC_i(v, j)$ to denote the expected cost of agent A_i, when in state (v, j), where v is the best value obtained after gathering information on j opportunities. The value of $EC_i(v, j)$ can be calculated recursively according to:

$$EC_i(v,j) = \begin{cases} 0 & r_i(j) \leq v < V_i^{IS} \vee r_i^{resume} \leq v \vee j = n_i \\ c_i + \int_{y=-\infty}^{\infty} EC_i(\max(v_0^i, y, v), j+1) f_i(y) dy & \text{otherwise} \end{cases} \tag{12}$$

The first case is where the agent unavoidably terminates its IGP (according to Theorem 2). The second case is where, the IGP resumes, hence the expected cost is given recursively based on the new state $(\max(v, y), j + 1)$ the agent will be in after gathering information on one additional opportunity. This allows us to calculate $EC_i(x)$, as: $EC_i(x) = EC_i^c(-\infty, 0)$.

These enable us to calculate the expected benefit of agent A_i when the other agents use the set of strategies $\{(p_i, (r_i(1), .., r_i(n_i)), r_i^{resume}) | 1 \leq i \leq k \wedge i \neq j\}$. If agent A_i will choose to engage in IGP then its expected benefit, $EB_i(IGP)$, is given by:

$$\begin{aligned} EB_i(IGP) = -EC_i + \int_{y=V_i^{IS}}^{\infty} \max(v_0^i, y) f_i^{\text{return}}(y) dy \\ + \int_{y=-\infty}^{V_i^{IS}} f_i^{\text{return}}(y) \int_{x=-\infty}^{\infty} \max(v_0^i, y, x) \bar{f}_i(x) dx dy \end{aligned} \tag{13}$$

which is similar to (4), except that the differentiation between the second and third terms on the right is according to V_i^{IS} rather than P_i^{IS}.

When the agent opts not to gather information at all, its expected benefit, $EB_i(\neg IGP)$, is simply the expected value of the maximum value returned by the other agents taking part in the ISP:

$$EB_i(\neg IGP) = \int_{y=-\infty}^{\infty} \max(y, v_0^i) \bar{f}_i(y) dy \tag{14}$$

The equilibrium stability conditions remain as in Section 3.1, replacing the calculation of $EB_i($ $IGP)$ and $EB_i(\neg IGP)$ with (13) and (14). As with the former method, there is no guarantee that an equilibrium will actually exist (either pure or mixed) and that if one exists there will be no other equilibria. Also, as with the method presented above, the analysis of the Threshold restricted method generalizes the analysis of the standard cooperative information gathering model in the

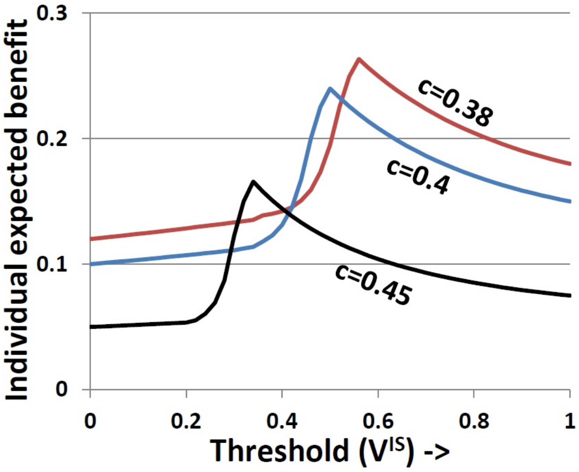

Figure 3: Threshold restricted information sharing - effect of V^{IS} on the individual expected benefit, for different information gathering costs, c, in a setting: $k = 2$ and $n = 1$.

sense that the latter is a specific case where the threshold set for taking part in the ISP is $V_i^{IS} \to \infty$ $\forall i$. Similarly, when $V_i^{IS} \to -\infty$ $\forall i$, the solution obtained is the same as the one known for the single-agent information gathering problem (McMillan & Rothschild, 1994).

Figure 3 illustrates the agents' individual expected benefit as a function of the thresholds V_i^{IS} used, for different information gathering costs (c). The setting used is the homogeneous setting described at the beginning of the section, with the parameters $c = 0.4$ and $n = 1$. The values V_i^{IS} were similar for all agents (i.e., $V_i^{IS} = V^{IS}$ $\forall i$). As depicted in the figure, the maximum expected individual benefit is obtained when the value of the threshold for excluding agents from the ISP is, in some cases, substantially small, leaving many favorable findings outside the ISP. The typical pattern exhibited in the figure is similar to the one depicted in Figure 1: an increase and then a decrease in the expected individual benefit as the threshold V^{IS} increases. This is explained by the fact that when $V^{IS} = 0$ each agent actually executes an individual IGP without sharing information with others. As V^{IS} increases, the agent relies more on the information gathered by the other agents and the same considerations explained in 3.1 regarding the tradeoff between the negative effect over p_i and r_i and the positive effect over the chance of actually taking part in the ISP hold.

3.3 Cost Filtered Information Sharing

This method introduces a cost c^{IS} an agent incurs if it chooses to take part in the ISP. For example, in our travel agents running example this is equivalent to having all agents that are interested in taking part in the ISP present some evidence of donating a fixed amount of money to some charity, as a prerequisite for accessing the designated secured server that is used for sharing. Unlike the two prior methods, here the agents get to decide whether to opt-out of information sharing; hence it does not require any enforcement whatsoever. It does require, however, a means for introducing a cost for the ISP, e.g., through a donation or by setting a meeting place for sharing information such that each participant will need to spend time and money for getting there.

It is assumed that all agents are a priori acquainted with the ISP participation costs of all agents. The introduction of such a cost (that is not eventually returned to the agents) requires an appropriate balance in the form of some compensation to the agents for taking part in the ISP. Without such compensation, no agent will be willing to take part in the ISP, as proven in the following proposition.

Proposition 1. *If the agents incur some cost when taking part in the ISP, then in the absence of some compensation for taking part in the ISP, none of the agents will take part in the ISP.*

Proof. Consider the highest value v that if known to any of the agents, after completing their individual IGP, warrants the participation of the agent in the ISP. We show that in the absence of an appropriate compensation for the agent, such a value v cannot hold — since none of the other agents will contribute a value greater than v to the ISP, the agent that found v will not gain anything from the ISP however it will incur a cost; consequently it will choose not to take part in the ISP. □

One option to compensate the agents for taking part in the ISP is to offer the agent with the best value that takes part in the ISP a compensation B. The amount B can be collected from the agents (e.g., in equal shares) prior to the IGP, such that once collected it is considered a "sunk cost" and the agents' strategies become affected only by the chance of receiving the bonus B. The structure of the best response strategy of any individual agent in this case, given the strategy of others, is given in Theorem 3.

Theorem 3. *Given the probability distribution function of the maximal value obtained by all other agents that take part in the ISP, $\bar{f}_i(x)$, agent A_i's optimal individual information gathering strategy can be described by the pair (r_i, R_i^{IS}), where R_i^{IS} is the set of intervals such that for any $x \in R_i^{IS} \wedge x \geq v_0^i$:*

$$c^{IS} \leq \int_{y=x}^{\infty} (y-x)\bar{f}_i(y)dy + B \cdot \int_{y=-\infty}^{x} \bar{f}_i(y)dy \tag{15}$$

and is to set a value $r_i \geq v_0^i$, where r_i is the solution to:

$$c_i = \int_{y=r_i}^{\infty} (EB_i(y) - EB_i(r_i)) f_i(y)dy \tag{16}$$

where $EB_i(v)$ is given by:

$$EB_i(v) = \begin{cases} v & v \notin R_i^{IS} \\ -c^{IS} + B \cdot \int_{y=-\infty}^{v} \bar{f}_i(y)dy & \textit{otherwise} \\ + \int_{y=-\infty}^{\infty} \max(v_0^i, y, v)\bar{f}_i(y)dy & \end{cases} \tag{17}$$

The agent should resume its individual IGP as long as the value found so far is below r_i, and otherwise it should terminate its IGP. Upon terminating the IGP (or obtaining the value of all opportunities) the agent should participate in the ISP if the best value it has found in its individual IGP is in one of the intervals of the set R_i^{IS} and otherwise it should opt out of taking part in the ISP.

Proof. The set R_i^{IS} as defined in (15) contains all the values v for which the expected benefit from taking part in the ISP - calculated as the potential value improvement, $\left(\int_{y=x}^{\infty}(y-x)\bar{f}_i(y)dy\right)$ plus the expected compensation $B \cdot \int_{y=-\infty}^{x} \bar{f}_i(y)dy$, both independent of the reservation value r_i used by the agent - is greater than the cost c^{IS} incurred. The remainder of the proof, concerning the optimality of a reservation-value-based strategy and the correctness of (16) is the same as the one provided for Theorem 1, differing only in the way the expected benefit if information gathering is resumed is calculated. □

The solution of a set of equations consisting of (16)-(17) for each agent will provide a set of pure equilibria of the form $\{(r_i, R_i^{IS})|1 \leq i \leq k\}$, if any exist. For the same considerations given in Section 3.1, a mixed equilibrium for this case will be of the form $\{(p_i, r_i, R_i^{IS})|1 \leq i \leq k\}$.

The equilibrium analysis of the Cost filtered information sharing follows the analysis given for the previous methods, therefore we provide only the differences. The calculation of $F_i^{\text{return}}(x)$ remains as in (2). The probability distribution function $f_i^{\text{return}}(x)$ is the first order derivative of $F_i^{\text{return}}(x)$. The function $F_i^{\text{return}'}(x)$, is now calculated as a modification of (10):

$$F_i^{\text{return}'}(x) = F_i^{\text{return}}(x) + \int_{y \geq x \wedge y \notin R_i^{IS}} f_i^{\text{return}}(y)dy$$

and consequently the function $\bar{F}_i(x)$ is given by:

$$\bar{F}_i(x) = \prod_{A_j \in K \wedge j \neq i} (p_j F_j^{\text{return}'}(x) + (1 - p_j))$$

The probability distribution function $\bar{f}_i(x)$ is the first order derivative of $\bar{F}_i(x)$ as before.

These enable us to calculate the expected benefit of agent A_i when the other agents use the set of strategies $\{(p_j, r_j, R_j^{IS})|1 \leq j \leq k \wedge i \neq j\}$. If agent A_i will choose to engage in IGP then its expected benefit, $EB_i(IGP)$, is given by:

$$EB_i(IGP) = -c_i \frac{1 - F_i(r_i)^{n_i}}{1 - F_i(r_i)} + \int_{y=-\infty}^{\infty} EB_i(\max(y, v_0^i)) f_i^{\text{return}}(y)dy \tag{18}$$

The first term is the expected cost incurred throughout the IGP carried out by A_i, calculated as: $c_i \sum_{j=1}^{n_i} (F_i(r_i))^{j-1} = c_i \frac{1-F_i(r_i)^{n_i}}{1-F_i(r_i)}$, as the number of opportunities on which information is gathered is a geometric random variable bounded by n_i, with a $1 - F_i(r_i)$ success probability. The second term is the expected benefit from executing its IGP. When the agent opts not to gather information at all, its expected benefit, $EB_i(\neg IGP)$, is simply the maximum between the expected value of the maximum value returned by the other agents and v_0^i, if the agent will choose to take part in the ISP:

$$EB_i(\neg IGP) = EB_i(v_0^i) \tag{19}$$

The equilibrium stability conditions remain as in Section 3.1, replacing the calculation of $EB_i(IGP)$ and $EB_i(\neg IGP)$ with (18) and (19). As with the former methods, there is no guarantee that an equilibrium will actually exist (either pure or mixed) and that if one exists there will be no other equilibria. Also, as with the methods presented above, the analysis of the cost filtered method generalizes the analysis of the standard cooperative information gathering model in the sense that the latter is a specific case where $c^{IS} = B = 0$. Similarly, when $B = 0$ the solution obtained is the

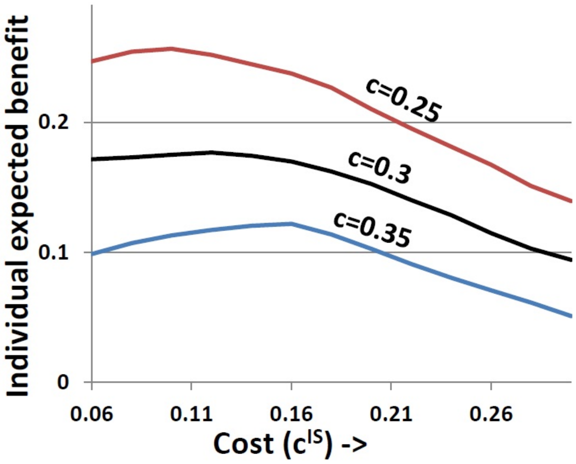

Figure 4: Cost filtered information sharing - effect of c^{IS} on the individual expected benefit, for different information gathering costs, c in a setting: $k = 2$, $n = 3$ and $B = 0.04$.

same as the one known for the single-agent information gathering problem (regardless of the value of c^{IS}, based on Proposition 1).

Figure 4 illustrates the agents' individual expected benefit as a function of the cost c^{IS} used, for different information gathering costs (c). The setting used is the homogeneous setting described at the beginning of the section, using the parameters $k = 2$, $n = 3$ and $B = 0.04$. As illustrated in the figure, the maximum expected individual benefit is obtained when a substantial cost is incurred for taking part in the ISP. The typical pattern exhibited in the figure is similar to the one depicted in Figures 1 and 3, and explained by similar considerations.

3.4 Random Finding Sharing

In this method each agent A_i, after it has completed its individual IGP, is allowed to take part in the ISP, however it is restricted to disclose only one of the values it has come across in its IGP, which is randomly selected. The downside of such a restriction is obvious - in most cases the agents will not benefit from the best individual findings but rather become exposed to a small, not necessarily optimal, subset of the information gathered. Still, as we demonstrate in the following paragraphs, for some settings the individual performance improves. In our travel agents running example the implementation of the method is quite straight forward and relies, as before, on having the agents send their findings to a designated secured server used for sharing. The only difference, however, is that the server will pick only a single random finding of each agent and will discard the rest of the findings. Here, again, no enforcement is necessary since no individual agent can benefit from disclosing more information than it is required to disclose.

The structure of the best response strategy of any individual agent in this case, given the strategy of others, is given in Theorem 4.

Theorem 4. *Given the probability distribution function of the maximal value obtained in the ISP based on the reports of all other agents, denoted $\bar{f}_i(x)$, agent A_i's optimal individual information* *gathering strategy is to follow a reservation value $r_i \geq v_0^i$, where r_i is the solution to:*

$$c_i = \int_{y=r_i}^{\infty} f_i(y) \int_{x=-\infty}^{\infty} (\max(y,x) - \max(r_i,x))\bar{f}_i(x)dxdy \tag{20}$$

The agent should always choose to gather information on an additional opportunity (if one is still available) if the best value obtained so far is below r_i and otherwise it should proceed to ISP.

Proof. The proof is similar to the one given for Theorem 1, differing only in the way the expected benefit if gathering information on an additional opportunity is calculated. □

The solution of a set of k equations of type (20), one for each agent, will provide a set of pure equilibria of the form $\{r_i|1 \leq i \leq k\}$ if any exist. For the same considerations given in Section 3.1, a mixed equilibrium for this case will be of the form $\{(p_i, r_i)|1 \leq i \leq k\}$.

The equilibrium analysis for this method follows the one given in 3.1, except that $F_i^{\text{return}'}(x)$ is used as in 3.2. The function $F_i^{\text{return}}(x)$ is given by (2). The probability distribution function $f_i^{\text{return}}(x)$ is the first order derivative of $F_i^{\text{return}}(x)$. In order to formulate $F_i^{\text{return}'}(x)$ we use a somewhat different state definition for $F_i^{\text{return}'}(x,v,j,l)$. A state (v,j,l) is now defined according to the number of values that are below x obtained by agent A_i after gathering information on j opportunities, denoted l, and the best value obtained so far (including v_0^i), v. The function $F_i^{\text{return}'}(x,v,j,l)$ can be calculated recursively according to:

$$F_i^{\text{return}'}(x,v,j,l) = \begin{cases} 0 & x < v_0^i \\ \frac{l}{j} & v \geq r_i \vee n_i = j \\ \int_{y=-\infty}^{x} F_i^{\text{return}'}(x, \max(y,v), j{+}1, l{+}1) f_i(y)dy & \text{otherwise} \\ \quad + \int_{y=x}^{\infty} F_i^{\text{return}'}(x, \max(y,v), j{+}1, l) f_i(y)dy & \end{cases} \tag{21}$$

The case where $x < v_0^i$ is trivial, as v_0^i is a lower bound for the best value the agent ends up with. The second case in (21) is straightforward — when the best value obtained so far is $v \geq r_i$, or when there are no additional opportunities ($n_i = j$), the agent necessarily terminates its IGP (according to Theorem 4). In such cases, if l values of the total j values are below x, then the probability that A_i will return a value below x is $\frac{l}{j}$. In all other cases, the IGP will resume; hence the probability is given recursively based on the new state $(x, \max(y,v), j+1, l+1)$ in case the new value obtained is $y \leq x$ and $(x, \max(y,v), j+1, l)$ otherwise.

Thus, we can now formulate $F_i^{\text{return}'}(x)$:

$$F_i^{\text{return}'}(x) = F_i^{\text{return}'}(x, -\infty, 0, 0) \tag{22}$$

and consequently $\bar{F}_i(x)$ is given by (11) and the probability distribution function $\bar{f}_i(x)$ is the first order derivative of $\bar{F}_i(x)$ as before.

These enable us to calculate the expected benefit of agent A_i when the other agents use the set of strategies $\{(p_j, r_j)|1 \leq j \leq k \wedge i \neq j\}$. If agent A_i will choose to engage in IGP then its expected benefit, $EB_i(IGP)$, is given by:

$$EB_i(IGP) = -\, c_i \frac{1 - F_i(r_i)^{n_i}}{1 - F_i(r_i)} + \int_{y=-\infty}^{\infty} f_i^{\text{return}}(y) \int_{x=-\infty}^{\infty} \max(v_0^i, y, x)\bar{f}_i(x)dxdy \tag{23}$$

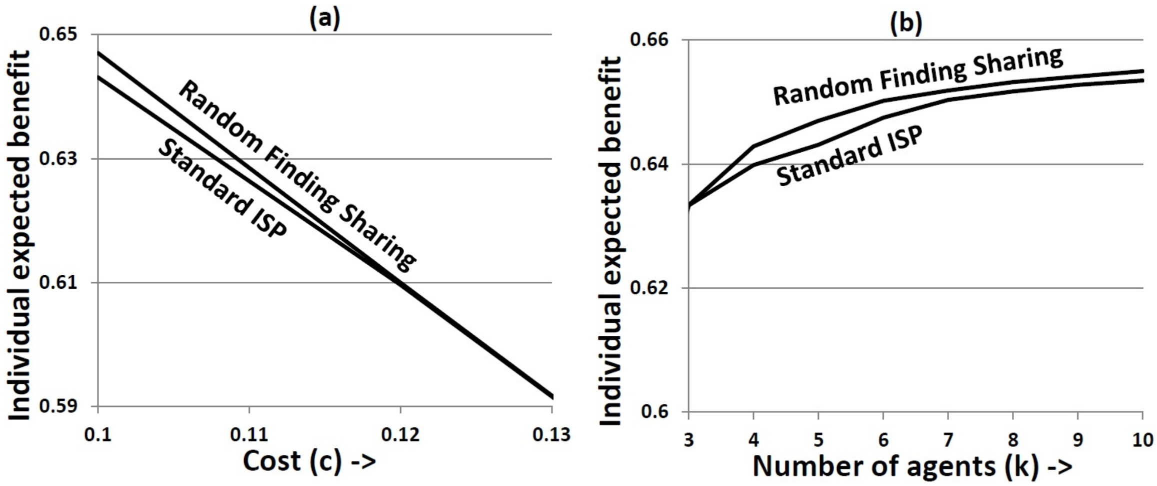

Figure 5: Random finding sharing and Standard information sharing (according to Rochlin et al., 2014) as a function of: (a) information gathering costs, c, in a setting: $k = 5$ and $n = 3$; and (b) number of agents, k, in a setting: $c = 0.1$ and $n = 3$.

When the agent opts not to gather information at all, its expected benefit, $EB_i(\neg IGP)$, is simply the expected value of the maximum value returned by the other agents:

$$EB_i(\neg IGP) = \int_{x=-\infty}^{\infty} \max(v_0^i, x)\bar{f}_i(x)dx \tag{24}$$

The equilibrium stability conditions remain as in Section 3.1, replacing the calculation of $EB_i(IGP)$ and $EB_i(\neg IGP)$ with (23) and (24). As with the former method, there is no guarantee that an equilibrium will actually exist (either pure or mixed) and that if one exists there will be no other equilibria.

Figure 5 depicts the agents' individual expected benefit as a function of information gathering costs c (left graph) and the number of agents k (right graph). The setting used is the homogeneous setting described in the beginning of the section, with the parameters $k = 5$ and $n = 3$ (Figure 5(a)) and $c = 0.1$ and $n = 3$ (Figure 5(b)). Each graph depicts the performance both with Random finding sharing and the standard non-restricted sharing method (according to Rochlin et al., 2014). The figure shows that the random finding sharing strategy dominates the standard information sharing strategy, as far as the expected individual benefit is concerned.

3.5 Subgroup Restricted Information Sharing

In this method the agents are divided into sub groups that share their findings separately, i.e., each agent executes its individual IGP separately and at the end each subgroup will carry out a separate ISP. Formally, consider the case where the $|K|$ agents are divided into w subgroups $\{K_1, \cdots, K_w\}$ ($\bigcap_{j=1}^{w} K_j = \emptyset$ and $\bigcup_{j=1}^{w} K_j = K$). In our travel agents running example this can be done by providing each sub-group a designated server that will serve only the subgroup members or, in physical environments, setting different meeting places for sharing the information after the group has been partitioned into sub-groups.

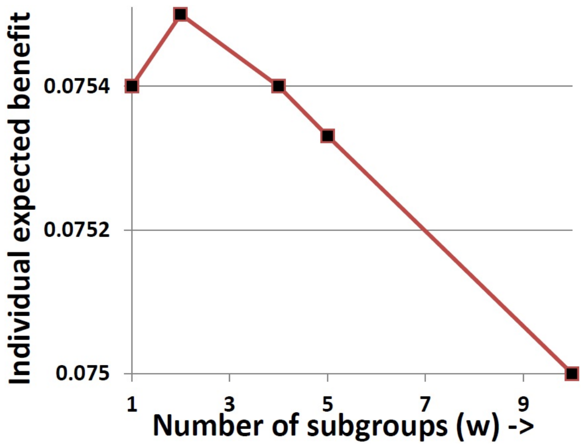

Figure 6: Subgroup restricted information sharing - individual expected benefit as a function of the number of subgroups, w, in a setting: $c = 0.45$, $k = 20$ and $n = 10$.

The best response strategy of any individual agent A_i^j of subgroup $K_j = \{A_1^j, \dots A_{|k_j|}^j\} \subseteq K$ as well as the equilibrium analysis within the subgroup level are the same as given by (Rochlin et al., 2014), as all agents within the subgroup fully share their findings. The optimal strategy can thus be obtained by checking the expected benefit of all possible divisions of K into subgroups and selecting the partition associated with the highest expected benefit. The computational complexity of evaluating all subgroups is combinatorial in the number of agents. Although the focus of the paper is not on the computational aspects but rather on analyzing the structure of the equilibrium cooperative strategies, we note that in most ISP settings the computational complexity becomes a non-issue since the number of agents taking part in the ISP is relatively small.

We note that this method results in a similar performance as in (Rochlin et al., 2014) when the number of subgroups is $w = 1$ (i.e., all k agents will take part in the ISP as one group). Similarly, when the number of subgroups is $w = |K|$ (i.e., each subgroup contains only one agent) the solution obtained is the same as the one known for the single-agent information gathering problem (McMillan & Rothschild, 1994).

Figure 6 illustrates the agents' individual expected benefit as a function of the number of subgroups used (w). The setting used is the homogeneous setting described in the beginning of the section, using the parameters $k = 20$, $n = 10$ and $c = 0.45$. The graph uses partitioning into equal-size subgroups, i.e., w obtains the values $\{1, 2, 4, 5, 10\}$. As depicted in the figure, the maximum expected individual benefit is obtained in this example when the number of subgroups is $w = 2$.

4. Related Work

The model analyzed in this paper is based on two important concepts: multi-agent cooperation and costly information gathering. Multi-agent cooperation has been shown to be widely effective for better achieving agents' individual goals (Stone & Kraus, 2010) or improve their performance measures (Kraus, Shehory, & Taase, 2003; Dutta & Sen, 2003), especially when there are differences in the agents' capabilities, knowledge and resources or when an agent is incapable of completing a task by itself (Stone & Kraus, 2010; Saad, Han, Debbah, & Hjorungnes, 2009; Conitzer, 2012; Breban & Vassileva, 2001). It is also the main driving force behind many coalition formation models in the area of cooperative game theory and MAS (Shehory & Kraus, 1998). Yet, the majority of cooperation and coalition formation MAS-related research tends to focus on the way coalitions are formed and consequently concerns issues such as the optimal division of agents into disjoint exhaustive coalitions, division of coalition payoffs and enforcement methods for interaction protocols. While coalition formation and coordination models can be widely found in the electronic market domain, most work in this domain emphasizes mechanisms for forming cooperation for the purpose of aggregating demands in order to obtain volume discounts (Tsvetovat, Sycara, Chen, & Ying, 2000; Yamamoto & Sycara, 2001; Sarne & Kraus, 2003). Several authors have considered the problem of determining the strategy of a group once formed (Ito, Ochi, & Shintani, 2002; Sarne, Manisterski, & Kraus, 2010; Rochlin, Sarne, & Zussman, 2011; Mash, Rochlin, & Sarne, 2012), however their focus was mostly on fully-cooperative agents. None of these works considered the cooperation problem of a group of self-interested agents in costly exploration settings where findings can benefit all agents.

Group-based cooperation of self-interested agents can also be found in public goods games and allocation games in general (Aumann, 1998; Nagel & Tang, 1998; McKelvey & Palfrey, 1992; de Jong et al., 2008; de Jong & Tuyls, 2011). Common to these games is that according to their equilibrium each agent individually should opt out of the cooperation as soon as possible or invest the minimum allowed. Therefore the research on cooperation in this domain is mostly studied for repeated games (Selten & Stoecker, 1986) or settings with bounded-rational participants (e.g., people) for which cooperation to some extent is commonly exhibited. Much effort has been placed on developing reciprocity-based mechanisms, e.g., tit-for-tat (Axelrod, 1984) that facilitate cooperation even when agents find it momentarily beneficial to act selfishly. This way, long-term considerations override short-term greedy behavior. Many have extended the basic mechanism to support various variants of the model, such as asymmetric costs, heterogeneously repeating instances and other factors (Sen, 1996). Few works has dealt with inducing cooperation in non-repeated settings showing that rewards are somewhat less effective than sanctions in enforcing cooperation (Walker & Halloran, 2004). The main difference between public goods games and our work is that in our case the "public goods" are in fact information that can lead to better economic decisions. Obtaining more information requires carrying out an active sequential information gathering process. Therefore, in our settings there is much room for individual information gathering, to some extent, even if all the others are "free riders". Moreover, with the simplistic settings used in public goods games the issue of information sharing and the ways it is carried out (when considering self-interested agents) becomes irrelevant.

The second concept upon which this paper relies, i.e., costly information gathering, is of great importance when there is no central source that can supply an agent full immediate reliable information on the environment and the state of the other agents (Sarne & Aumann, 2014). In general, the introduction of information gathering costs into MAS models leads to a more realistic description of these environments. This is because agents are typically required to invest/consume some of their resources in order to obtain information concerning opportunities available in their environment (Bakos, 1997; Sarne & Kraus, 2008; Kephart & Greenwald, 2002; Rochlin & Sarne, 2013; Rochlin, Sarne, & Zussman, 2013; Manisterski, Sarne, & Kraus, 2008).

Optimal strategies for settings where individuals need to search for an applicable opportunity when information gathering is costly have been widely studied (Grosfeld-Nir, Sarne, & Spiegler, 2009; Elmalech, Sarne, & Grosz, 2015; Elmalech & Sarne, 2012), prompting several literature reviews (Smith, 2011; McMillan & Rothschild, 1994; Morgan & Manning, 1985). These models, which are often termed "costly search" models or "economic search" models have been developed to the point where their total contribution is referred to as "search theory". Over the years, many information gathering model variants have been considered, focusing on different aspects of the model, such as the decision horizon (finite versus infinite) (Lippman & McCall, 1976), the presence of the recall option (McMillan & Rothschild, 1994), the option to disambiguate noisy signals (Chhabra, Das, & Sarne, 2014a; Alkoby, Sarne, & Das, 2015; Chhabra, Das, & Sarne, 2014b), the distribution of values and the extent to which findings remain valid along the process (Landsberger & Peled, 1977). In particular, these models have been integrated in the study of strategic information platforms (Hajaj, Hazon, Sarne, & Elmalech, 2013; Sarne, 2013; Hajaj & Sarne, 2014). Another strand of search-based models is two-sided search (Sarne & Arponen, 2007; Hendrix & Sarne, 2007) that deals with the distributed (search-based) formation of pairwise (or general size) partnerships (Nahum, Sarne, Das, & Shehory, 2015). While the analysis of such models derive from equilibrium considerations, these are very different from our model as they do not reflect any cooperative aspect.

Many cooperative information gathering models have been studied, extending the theories to multi-agent (or multi-goal) environments (Sarne & Kraus, 2005). Examples include, among others, an attempt to purchase several commodities while facing imperfect information concerning prices or operating several robots in order to evaluate opportunities in different locations. These works differ from ours either in that they consider fully cooperative agents that attempt to maximize the overall utility (Sarne et al., 2010; Gatti, 1999; Burdett & Malueg, 1981; Carlson & McAfee, 1984), and thus lack any equilibrium considerations, or they assume that any agent's IGP is constrained by the findings of the other agents, rather than augmented/improved by such findings as in our case (Rochlin, Sarne, & Laifenfeld, 2012; Rochlin & Sarne, 2014b). Consequently they constitute substantially different equilibrium strategies. Models that do consider cooperative information gathering, which rely on assumptions similar to ours (e.g., Rochlin et al., 2014; Hazon et al., 2013), focus primarily on the extraction of the equilibrium strategies and investigate the influence of the different model parameters on the agents' performance in equilibrium. None of these works, however, suggested methods for improving cooperative information gathering in such settings, of the kind that we suggest and analyze in this paper.

More broadly, our problem can be seen as part of the field of planning under uncertainty, hence it is related to Markov decision processes (MDP) (Bellman, 1957; Puterman, 1994) and decentralized Markov decision processes (Bernstein, Givan, Immerman, & Zilberstein, 2002). In these models the goal is to maximize the expected cumulative reward, which is also the objective in our case. Alas, the use of MDPs in our case is complicated by the continuous nature of the value probability distribution functions. More importantly, our analysis and proofs result in threshold-based (or interval-based) solutions which are both simpler in terms of strategy and state representation and can be derived with a substantially lesser complexity compared to solving via MDPs.

Finally, we note that the non-intuitive findings whereby methods that essentially limit information sharing and cooperation actually have a positive impact in the self-interested case follows, in spirit, earlier results in other settings. In particular, ones in which it has been shown that so-called "inefficiencies" can increase market performance, under certain circumstances. For example, Masters (1999) shows that an increase in minimum wage, which is often considered inefficiency in economics, can have positive employment effects. In transportation economics (e.g., congestion games) equilibrium is frequently not the overall optimum. In such cases, it has been shown that taxation can change the equilibrium to a more desirable one (Penn, Polukarov, & Tennenholtz, 2009b, 2009a; Fotakis, Karakostas, & Kolliopoulos, 2010). Similarly, taxes can facilitate more desirable equilibria in Boolean games (Endriss, Kraus, Lang, & Wooldridge, 2011) and in centralized matching schemes (Anshelevich, Das, & Naamad, 2013). In this work we show that a somewhat similar phenomenon also occurs in the context of cooperative information gathering, though the model and analysis are, of course, totally different from the above mentioned.

5. Discussion and Conclusions

As demonstrated in Section 3, each of the five methods proposed and analyzed in this paper can substantially increase the benefit self-interested agents achieve through information sharing when gathering information cooperatively. Each of the five methods is based on a different restriction made on the agents' ability or willingness to take part in the ISP. Intuitively such restrictions may seem to have a negative effect on performance. Yet, since each agent gains less from the information sharing itself, it has a greater incentive to invest more resources in individual information gathering, hence overall performance improves.

The results suggest important inputs for the designers of markets and systems where cooperative information gathering is applicable, by enabling them to predict the strategies that will be used and the resulting system performance. These primarily facilitate the proper design of the system and the determination of what elements should and should not be included in such systems in order to achieve specific goals and promote certain behavior. In particular, the introduction of some seemingly non-beneficial elements may actually be productive. We note that the paper generally does not attempt to find the "optimal" parameter values for each method (e.g., probability, group partitioning, threshold and cost of taking part in the information sharing, or the payment received if the agent is associated with the "best" value), since the concept of optimality in this sense is not properly defined. Indeed in settings where there is an equilibrium solution that is preferred by all agents (e.g., in the examples given in the former sections, where all agents are homogeneous) the choice of parameter values is clear. Nevertheless in general, it is possible that a certain value will be preferred by one of the agents whereas others will prefer another. In the latter case it is the role of the system designer to decide on these parameters based on her goals.

There are numerous extensions of the model that can be considered. Some of them are straightforward and require minor changes in the analysis. For example, if agents are buyers and each of them is interested in more than a single unit of the product they are searching for, the only change required in the individual strategy equations is multiplication of the expense of purchasing the item by the number of items in which the agent is interested. Another example is the composition of several methods. For example, in the case of Subgroup restricted information sharing, each of the subgroups can adopt any of the other four methods in the subgroup level. Other extensions, while of much interest, are more complex to analyze. For example, consider a model where the agents can continuously share their findings along their individual IGPs. In this case, as discussed in Section 1, it is essential to first define the method that will provide an incentive for agents to share their findings despite the negative influence it will have in terms of discouraging others from further information gathering.

Acknowledgments

Preliminary results of this work appeared in the Proceedings of the Thirteenth International Conference on Autonomous Agents and Multiagent Systems (Rochlin & Sarne, 2014a). The authors would like to thank Barbara Grosz for insightful comments that helped improve this paper substantially. The first coauthor was a student at Bar-Ilan University when the research reported in this paper was carried out. This research was partially supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1083/13) and the ISF-NSFC joint research program (grant No. 2240/15).

Appendix A. Proof of Theorem 1

Proof. We first prove the reservation-value nature of the optimal strategy. Then we continue with an inductive proof, to show that the reservation value used by each agent will remain stationary along its IGP and can be calculated according to Equation 1.

In the absence of any other new information along the IGP, each agent's strategy will be the mapping of $S(\vec{x}, j) \rightarrow \{terminate, resume\}$, where $\vec{x}$ is the set of values obtained so far (including v_0^i) and j is the number of opportunities on which information has already been gathered. Since the agent is interested merely in the maximum opportunity value, its strategy is affected only by the maximum value in $\vec{x}$, hence the strategy can be defined as $S(v, j) \rightarrow \{terminate, resume\}$, where v is the maximum value in $\vec{x}$. Obviously, if according to the optimal strategy the agent needs to resume IGP upon reaching state (v, j) then the same should be true for any state (v', j) where $v' < v$. Similarly, if according to the optimal strategy the exploration should terminate at state (v, j) then the same should hold for any state (v'', j) where $v'' > v$. Therefore, for each given number of opportunities for which information has been gathered, j, the optimal individual IGP strategy of agent A_i can be characterized by the reservation value r_i^j such that the agent should resume IGP if the best value obtained so far is below r_i^j and otherwise terminate the IGP.

We begin with the case of $j = n_i - 1$. If the best value obtained thus far by agent A_i is v then gathering information on one last opportunity, according to this strategy, will incur a cost c_i and the expected value the agents will obtain will be:

$$P_i^{IS} \cdot \int_{y=-\infty}^{\infty} f_i(y) \int_{z=-\infty}^{\infty} (\max(y, v, z)\bar{f}_i(z))dzdy \\ + (1 - P_i^{IS}) \cdot \int_{y=-\infty}^{\infty} \max(y, v) f_i(y)dy$$

The first term relates to the case where agent A_i will participate in the ISP, with a probability of P_i^{IS}, whereas the second term relates to the case where it will be required to opt out from the ISP, i.e., with a probability of $1 - P_i^{IS}$ (where y is the new value that will be obtained by the agent and z is the best value that will obtained from the other agents' IGP).

On the other hand, terminating the IGP at this point will result in a benefit:

$$P_i^{IS} \cdot \int_{z=-\infty}^{\infty} \max(v, z)\bar{f}_i(z)dz + (1 - P_i^{IS}) \cdot v$$

Therefore, the agent should gather information on the last opportunity if and only if:

$$
\begin{aligned}
&P_i^{IS} \cdot \int_{z=-\infty}^{\infty} \max(v,z)\bar{f}_i(z)dz + (1-P_i^{IS}) \cdot v < \\
&P_i^{IS} \cdot \int_{y=-\infty}^{\infty} f_i(y) \int_{z=-\infty}^{\infty} (\max(y,v,z)\bar{f}_i(z))dzdy \\
&+ (1-P_i^{IS}) \cdot \int_{y=-\infty}^{\infty} \max(y,v) f_i(y)dy - c_i
\end{aligned}
\tag{25}
$$

The left hand side of the equation captures the expected benefit if the individual IGP is terminated and the right hand side captures the expected benefit if information is gathered for the last opportunity. Both terms distinguish between the case where agent A_i participates in the ISP i.e., with a probability of P_i^{IS}, and when it is not allowed to. Using simple mathematical manipulations obtains:

$$
\begin{aligned}
&0 < P_i^{IS} \cdot \int_{y=v}^{\infty} f_i(y) \int_{z=-\infty}^{\infty} (\max(y,z) - \max(v,z)\bar{f}_i(z))dzdy \\
&+ (1-P_i^{IS}) \cdot \int_{y=v}^{\infty} (y-v) f_i(y)dy - c_i
\end{aligned}
\tag{26}
$$

The value v for which (26) becomes an equality is in fact the value of r_i according to (1). Therefore, since the right hand side of (26) is a decreasing function of v, then the agents should gather information on the last opportunity whenever the value of v is less than the value of r according to (1). This establishes the first part of the proof.

Now assume the same r_i (according to (1)) holds for any $j' > j$, for some j, and consider the agents' decision regarding gathering information on one more opportunity, if the best value obtained thus far is v and the number of opportunities for which the values were already obtained is j. If $v > r_i \geq v_0^i$ and the agent gathers information on one additional opportunity, then regardless of the value obtained next it will definitely terminate its individual IGP thereafter (as it already has a value greater than r_i and according to the induction assumption the optimal strategy thereafter is the reservation value r_i). Therefore the benefit obtained from further information gathering is given by:

$$
\begin{aligned}
&P_i^{IS} \cdot \int_{y=-\infty}^{\infty} f_i(y) \int_{z=-\infty}^{\infty} (\max(y,z) - \max(v,z)\bar{f}_i(z))dzdy \\
&+ (1-P_i^{IS}) \cdot \int_{y=-\infty}^{\infty} (y-v) f_i(y)dy - c_i
\end{aligned}
$$

Alas, since the latter term decreases as v increases, and obtains zero for $v = r_i$ (according to (1)), then since $v > r_i \geq v_0^i$ the term obtains a negative value, hence additional information gathering cannot be the preferred choice.

Similarly, consider the case where $v_0^i \leq v < r_i$ for j and the agent chooses not to gather additional information. Here the expected benefit from resuming information gathering is necessarily greater than if resuming in state $(v, j' > j)$. However, according to the induction assumption the agent should resume information gathering in state $(v, j' > j)$, leading to a contradiction. Therefore, the optimal strategy for j is also a reservation value strategy and the optimal reservation value is calculated, once again, according to (1). □

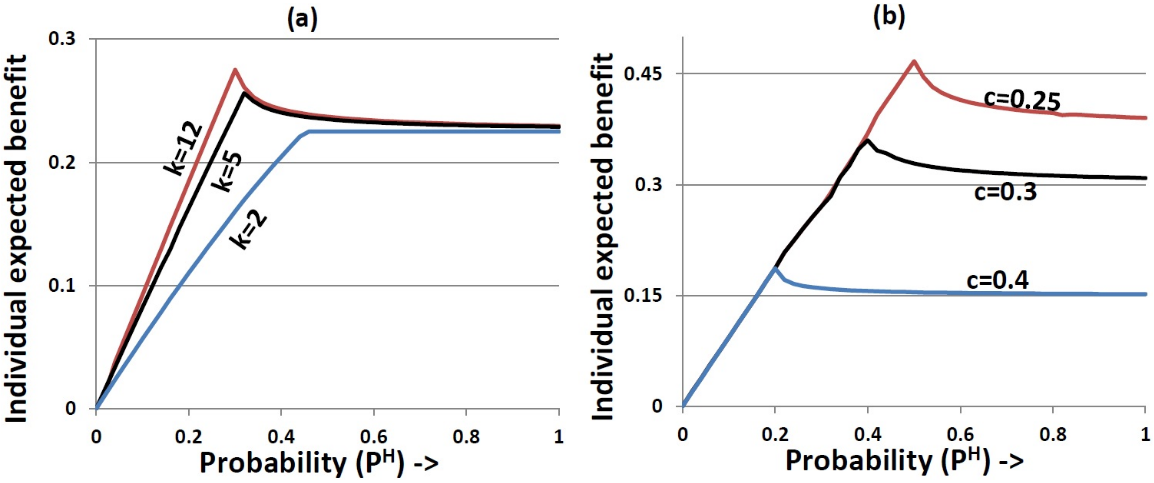

Figure 7: Enforced probabilistic information sharing - effect of P^H on the individual expected benefit, for different: (a) numbers of agents, k, in a setting: $c = 0.35$ and $n = 5$; and (b) information gathering costs, c, in a setting: $k = 15$ and $n = 4$.

Appendix B. Partially Limiting Information Sharing

Consider the case where each agent A_i, after it has completed its individual IGP, shares its own findings however can only obtain the information of others with a probability of P_i^H, which is a priori assigned. The structure of the best response strategy of any individual agent in this case, given the strategy of others, is identical to the one given in Theorem 1, replacing P_i^{IS} with P_i^H. Similarly the rest of the analysis given above hold for this case, except for the calculation of $\bar{F}_i(x)$, which is now given by:

$$\bar{F}_i(x) = \prod_{A_j \in K \wedge j \neq i} (p_j F_j^{\text{return}}(x) + (1 - p_j)) \tag{27}$$

Figure 7 depicts the agents' individual expected benefit as a function of the probability P_i^H used, for different group sizes (k) and information gathering costs (c) using the same setting used in Figure 1. A comparison of Figures 1 and 7 reveals that in this case the partial limitation over the information sharing dominates the enforced probabilistic information sharing strategy, as far as the expected individual benefit is concerned. Nonetheless, this is certainly not a general result.

Appendix C. Proof of Theorem 2

Proof. The proof generally resembles the one given for Theorem 1. If the best value agent A_i has found so far (including v_0^i) is $v > V_i^{IS}$ it is not allowed to take part in the ISP. If $j = n_i - 1$ then resuming the IGP will result in expected benefit $-c_i + \int_{y=-\infty}^{\infty} \max(y, v) f_i(y) dy$ whereas terminating the IGP will guarantee v. The agent should explore further for any v for which the first term is greater than the latter, which can be represented, after some mathematical manipulations, as:

$$c_i < \int_{y=v}^{\infty} (y - v) f_i(y) dy$$

which is equivalent to any value v greater than r_i^{resume} according to 6.[8]

The remainder of the proof which deals with showing the reservation-value nature of the optimal strategy and that if $v > V_i^{IS}$ the reservation value used by each agent remains stationary along its IGP is the same as in Theorem 1.

Now consider the case where the best value agent A_i has found so far is $v_0^i \leq v \leq V_i^{IS}$ and the number of opportunities on which information has already been gathered is j. Resuming the IGP will result in:

$$-c_i + \int_{y=-\infty}^{V_i^{IS}} f_i(y) \int_{z=-\infty}^{\infty} (\max(y, v, z) \bar{f}_i(z)) dz dy + \int_{y=V_i^{IS}}^{\infty} EB_i^{j+1}(y) f_i(y) dy$$

where $EB_i^{j+1}(y)$ is the expected benefit of agent A_i, given the best value it has obtained thus far, y (including v_0^i), and the number of opportunities on which information has already been gathered, j. The term $EB_i^j(v)$ is calculated recursively according to (8). If there are no remaining opportunities ($j > n$) or the best value found is $v \geq r_i^{resume}$ the expected benefit is simply v. Otherwise, the IGP resumes (incurring a cost c_i) with the revised best value ($\max(y, v)$), however with $j + 1$ opportunities on which information has already been gathered, i.e., with an expected benefit of $EB_i^{j+1}(\max(y, v))$. On the other hand, terminating the IGP at this point will result in the expected benefit:

$$\int_{z=-\infty}^{\infty} \max(v, z) \bar{f}_i(z) dz$$

Therefore, the agent should explore further for any v for which the first term is greater than the second, which after some mathematical manipulations becomes:

$$\begin{aligned} c_i < & \int_{y=v}^{V_i^{IS}} f_i(y) \int_{z=-\infty}^{\infty} (\max(y, z) - \max(v, z)) \bar{f}_i(z) dz dy \\ & + \int_{y=V_i^{IS}}^{\infty} f_i(y) \int_{z=-\infty}^{\infty} (EB_i^{j+1}(y) - \max(v, z)) \bar{f}_i(z) dz dy \end{aligned}$$

which is equivalent to any value greater than $r_i(j)$ according to 7. The reason a different threshold is used for different j values is that the calculation of the threshold depends among others on the expected benefit in case the value to be obtained is $v > V_i^{IS}$, in which case the expected benefit depends on the number of remaining opportunities. □

8. If the value of r_i^{resume} resulting from $c_i = \int_{y=v}^{\infty} (y-v) f_i(y) dy$ is lower than V_i^{IS}, the agent will inevitably terminate the IGP (as $v > r_i^{resume}$) hence any value $r_i^{resume} \leq V_i^{IS}$ can be used, in particular $r_i^{resume} = V_i^{IS}$ as in the theorem.

References

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

Anshelevich, E., Das, S., & Naamad, Y. (2013). Anarchy, stability, and utopia: creating better matchings. *Autonomous Agents and Multi-Agent Systems*, *26*(1), 120–140.

Aumann, R. (1998). On the centipede game. *Games and Economic Behavior*, *23*(1), 97–105.

Axelrod, R. (1984). *The evolution of cooperation*. Basic Books.

Bakos, Y. (1997). Reducing buyer search costs: Implications for electronic marketplaces. *Management Science*, *42*, 1676–1692.

Bellman, R. (1957). A Markovian decision process. *Indiana University Mathematics Journal*, *6*(4), 679–684.

Bernstein, D., Givan, D., Immerman, N., & Zilberstein, S. (2002). The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research*, *27*(4), 819–840.

Breban, S., & Vassileva, J. (2001). Long-term coalitions for the electronic marketplace. In *Proceedings of the E-Commerce Applications Workshop, Canadian AI Conference*.

Burdett, K., & Malueg, D. A. (1981). The theory of search for several goods. *Journal of Economic Theory*, *24*(3), 362–376.

Carlson, J. A., & McAfee, R. P. (1984). Joint search for several goods. *Journal of Economic Theory*, *32*(2), 337–345.

Chhabra, M., Das, S., & Sarne, D. (2014a). Competitive information provision in sequential search markets. In *Proceedings of the International conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2014)*, pp. 565–572.

Chhabra, M., Das, S., & Sarne, D. (2014b). Expert-mediated sequential search. *European Journal of Operational Research*, *234*(3), 861–873.

Chhabra, M., & Das, S. (2011). Learning the demand curve in posted-price digital goods auctions. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pp. 63–70.

Conitzer, V. (2012). Computing game-theoretic solutions and applications to security. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, pp. 2106–2112.

de Jong, S., & Tuyls, K. (2011). Human-inspired computational fairness. *Autonomous Agents and Multi-Agent Systems*, *22*(1), 103–126.

de Jong, S., Tuyls, K., & Verbeeck, K. (2008). Fairness in multi-agent systems. *Knowledge Engineering Review*, *23*(2), 153–180.

Dutta, P., & Sen, S. (2003). Forming stable partnerships. *Cognitive Systems Research*, *4*(3), 211–221.

Elmalech, A., Sarne, D., & Grosz, B. J. (2015). Problem restructuring for better decision making in recurring decision situations. *Autonomous Agents and Multi-Agent Systems*, *29*(1), 1–39.

Elmalech, A., & Sarne, D. (2012). Evaluating the applicability of peer-designed agents in mechanisms evaluation. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 02*, pp. 374–381.

Endriss, U., Kraus, S., Lang, J., & Wooldridge, M. (2011). Designing incentives for boolean games. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pp. 79–86.

Fotakis, D., Karakostas, G., & Kolliopoulos, S. G. (2010). On the existence of optimal taxes for network congestion games with heterogeneous users. In *Proceedings of the Third international conference on Algorithmic game theory (SAGT'10)*, pp. 162–173.

Gatti, J. (1999). Multi-commodity consumer search. *Journal of Economic Theory*, *86*(2), 219–244.

Grosfeld-Nir, A., Sarne, D., & Spiegler, I. (2009). Modeling the search for the least costly opportunity. *European Journal of Operational Research*, *197*(2), 667–674.

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

Hajaj, C., & Sarne, D. (2014). Strategic information platforms: selective disclosure and the price of "free". In *ACM Conference on Economics and Computation (EC'14)*, pp. 839–856.

Hazon, N., Aumann, Y., Kraus, S., & Sarne, D. (2013). Physical search problems with probabilistic knowledge. *Artificial Intelligence*, *196*, 26–52.

Hendrix, P., & Sarne, D. (2007). The effect of mediated partnerships in two-sided economic search. In Klusch, M., Hindriks, K., Papazoglou, M., & Sterling, L. (Eds.), *Cooperative Information Agents XI (CIA 2007)*, Vol. 4676 of *Lecture Notes in Computer Science*, pp. 224–240. Springer Berlin Heidelberg.

Ito, T., Ochi, H., & Shintani, T. (2002). A group-buy protocol based on coalition formation for agent-mediated e-commerce. *International Journal of Computing and Information Sciences*, *3*(1), 11–20.

Janssen, M. C. W., Moraga-Gonzalez, J. L., & Wildenbeest, M. R. (2005). Truly costly sequential search and oligopolistic pricing. *International Journal of Industrial Organization*, *23*(5-6), 451–466.

Kephart, J., & Greenwald, A. (2002). Shopbot economics. *Journal of Autonomous Agents and Multi-Agent Systems*, *5*(3), 255–287.

Kraus, S., Shehory, O., & Taase, G. (2003). Coalition formation with uncertain heterogeneous information. In *Proceedings of the Second International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2003)*, pp. 1–8.

Landsberger, M., & Peled, D. (1977). Duration of offers, price structure, and the gain from search. *Journal of Economic Theory*, *16*(1), 17–37.

Lippman, S., & McCall, J. (1976). The economics of job search: A survey. *Economic Inquiry*, *14*(3), 347–368.

Manisterski, E., Sarne, D., & Kraus, S. (2008). Enhancing cooperative search with concurrent interactions. *Journal of Artificial Intelligence Research (JAIR)*, *32*(1), 1–36.

Mash, M., Rochlin, I., & Sarne, D. (2012). Join me with the weakest partner, please. In *Proceedings of the 2012 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2012)*, pp. 17–24.

Masters, A. M. (1999). Wage posting in two-sided search and the minimum wage. *International Economic Review*, *40*(4), 809–826.

McKelvey, R., & Palfrey, T. (1992). An experimental study of the centipede game. *Econometrica*, *60*(4), 803–836.

McMillan, J., & Rothschild, M. (1994). Search. In *Proceedings of Handbook of Game Theory with Economic Applications*, pp. 905–927.

Morgan, P. (1983). Search and optimal sample sizes. *Review of Economic Studies*, *50*(4), 659–675.

Morgan, P., & Manning, R. (1985). Optimal search. *Econometrica*, *53*(4), 923–944.

Nagel, R., & Tang, F. (1998). Experimental results on the centipede game in normal form: an investigation on learning. *Journal of Mathematical Psychology*, *42*(2-3), 356–384.

Nahum, Y., Sarne, D., Das, S., & Shehory, O. (2015). Two-sided search with experts. *Autonomous Agents and Multi-Agent Systems*, *29*(3), 364–401.

Penn, M., Polukarov, M., & Tennenholtz, M. (2009a). Random order congestion games. *Mathematics of Operations Research*, *34*(3), 706–725.

Penn, M., Polukarov, M., & Tennenholtz, M. (2009b). Taxed congestion games with failures. *Annals of Mathematics and Artificial Intelligence*, *56*(2), 133–151.

Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience.

Rochlin, I., Aumann, Y., Sarne, D., & Golosman, L. (2014). Efficiency and fairness in team search with self-interested agents. In *Proceedings of the International conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2014)*, pp. 365–372.

Rochlin, I., & Sarne, D. (2014a). Constraining information sharing to improve cooperative information gathering. In *Proceedings of the 13th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*, pp. 237–244.

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

Rochlin, I., Sarne, D., & Laifenfeld, M. (2012). Coordinated exploration with a shared goal in costly environments. In *Proceedings of the ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track*, pp. 690–695.

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

Rochlin, I., Sarne, D., & Zussman, G. (2011). Sequential multilateral search for a common goal. In *Proceedings of the 2011 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2011)*, pp. 349–356.

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

Rochlin, I., & Sarne, D. (2013). Information sharing under costly communication in joint exploration. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, pp. 847–853.

Rothschild, M. (1974). Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, *82*(4), 689–711.

Saad, W., Han, Z., Debbah, M., & Hjorungnes, A. (2009). Coalitional games for distributed collaborative spectrum sensing in cognitive radio networks. In *Proceedings of the IEEE INFOCOM*, pp. 2114–2122.

Sarne, D. (2013). Competitive shopbots-mediated markets. *ACM Transactions on Economics and Computation*, *1*(3), 17.

Sarne, D., & Arponen, T. (2007). Sequential decision making in parallel two-sided economic search. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, p. 69.

Sarne, D., & Aumann, Y. (2014). Exploration costs as a means for improving performance in multiagent systems. *Annals of Mathematics and Artificial Intelligence*, *72*(3-4), 297–329.

Sarne, D., & Kraus, S. (2005). Cooperative exploration in the electronic marketplace. In *Proceedings of the Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference (AAAI 2005)*, pp. 158–163.

Sarne, D., Manisterski, E., & Kraus, S. (2010). Multi-goal economic search using dynamic search structures. *Autonomous Agents and Multi-Agent Systems*, *21*(1-2), 204–236.

Sarne, D., & Kraus, S. (2003). The search for coalition formation in costly environments. In *Proceedings of the Cooperative Information Agents VII, 7th International Workshop, CIA*, pp. 117–136.

Sarne, D., & Kraus, S. (2008). Managing parallel inquiries in agents two-sided search. *Artificial Intelligence*, *172*(4-5), 541–569.

Selten, R., & Stoecker, R. (1986). End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach. *Journal of Economic Behavior & Organization*, *7*(1), 47–70.

Sen, S. (1996). Reciprocity: a foundational principle for promoting cooperative behavior among self-interested agents. In *Proceedings of the Second International Conference on Multi-Agent Systems*, pp. 322–329.

Shehory, O., & Kraus, S. (1998). Methods for task allocation via agent coalition formation. *Artificial Intelligence*, *101*(1-2), 165–200.

Smith, L. (2011). Frictional matching models. *Annual Reviews in Economics*, *3*(1), 319–338.

Stone, P., & Kraus, S. (2010). To teach or not to teach? decision making under uncertainty in ad hoc teams. In *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pp. 117–124.

Tang, Z., Smith, M. D., & Montgomery, A. (2010). The impact of shopbot use on prices and price dispersion: Evidence from online book retailing. *International Journal of Industrial Organization*, *28*(6), 579–590.

Tsvetovat, M., Sycara, K., Chen, Y., & Ying, J. (2000). Customer coalitions in electronic markets. In *Proceedings of Agent-Mediated Electronic Commerce III, Current Issues in Agent-Based Electronic Commerce Systems (includes revised papers from AMEC 2000 Workshop)*, pp. 121–138.

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

Walker, J., & Halloran, M. (2004). Rewards and sanctions and the provision of public goods in one-shot settings. *Experimental Economics*, *7*(3), 235–247.

Yamamoto, J., & Sycara, K. (2001). A stable and efficient buyer coalition formation scheme for e-marketplaces. In *Proceedings of the 5th international conference on Autonomous agents (AGENTS '01)*, pp. 576–583.