Identifying the Best Agent in a Network

Leonie Baumann[*][†]

20th April 2017

[*]Cambridge INET-Institute, Faculty of Economics, University of Cambridge, United Kingdom. Email: lb714@cam.ac.uk.

[†]I am especially grateful to Anke Gerber and Ariel Rubinstein for comments and discussion. I also thank Sanjeev Goyal, Mikhail Safronov, participants in the Networks Reading Group at the Cambridge INET-Institute, and seminar participants at the University of Málaga.

Abstract

This paper develops a mechanism for a principal to assign a prize to the most valuable agent from a set of heterogeneously valued agents on a network. The principal does not know the value of any agent. Agents are competing for the prize and they have a „knowledge network“: If two agents are linked, they know each other's value for the principal. Each agent sends a costless private message to the principal about her own value and the values of other agents she knows. This message can be truthful or not. However, it is common knowledge that agents can lie about each value only to a certain extent and that agents only lie if lying increases their chances of being selected, otherwise they report truthfully. A mechanism which determines the probability of getting the prize for each agent for each possible message profile is proposed. It is shown that if every agent is linked to at least one other agent in the knowledge network, then the mechanism ensures existence of an equilibrium in which the most valuable agent gets the prize with certainty. If the knowledge network is complete or is a star, then the mechanism ensures that in every equilibrium the most valuable agent gets the prize with certainty.

Keywords: network, knowledge network, information network, mechanism design, principal-agents problem, allocation problem

JEL Classification Codes: C72, D82, D83

1 Introduction

This paper considers a mechanism design problem in a network setting: A principal has to assign a prize to one agent from a set of heterogeneously valued agents. She does not know the exact value of any agent and she wishes to assign the prize to the most valuable. Every agent would like to get the prize and knows her own value for the principal. So far, this is a standard allocation problem in the mechanism design literature which has many real-world-interpretations: A government wants to privatize a public service and has to select one firm out of many applying firms to provide the service from then on; an employer has to select one applicant out of all applicants for a position; a committee has to assign the prize to one individual among all nominees. Agents might have knowledge about the values of other applicants: two firms operating in the same field might have a good estimate about each others' costs; two people who have worked on a joint project might know each others' abilities. Such an information structure among agents can be interpreted as a *knowledge network* in which links between agents signify knowledge. The general concern of this paper is to propose a mechanism for the principal to identify the highest value agent with certainty in such a setting.

In the model introduced in this paper, agents exactly know their own value and the values of agents to whom they are linked in the knowledge network. We restrict our attention to knowledge networks in which every agent is linked to at least one other agent. The principal does not know any exact value, but the distribution of values and the network are common knowledge. Every agent sends a private costless message about her information to the principal. An agent's message contains a statement about her own value (*application*) and statements about the value of each of the agents she is linked to (*references*). Agents can lie about a true value, but only to a commonly known certain extent. This accounts for the intuition that statements must be to some extent credible and verifiable. For example, a scholar with a PhD in Economics cannot pass off as a PhD in Biology as her degree certificate specifies Economics as a field. However, it is possible to lie about her specialization within the field of Economics. Likewise, a junior applicant cannot pass off as a senior, but it is possible to lie about certain abilities of the junior.

Before agents send their messages, the principal publicly announces a mechanism according to which she will select an agent for the prize given a profile of messages. This mechanism defines a probability of being selected for every agent for every message profile. Knowing the mechanism, agents engage in a static Bayesian game. They simultaneously choose and send their messages. Each agent selects her message such that it maximizes her expected probability of being selected given her own type and all other agents' strategies. We introduce a *truth-telling assumption*: If the true message maximizes an agent's expected probability of being selected, then she strictly prefers to tell the truth over lying. The solution concept applied to this static game among agents is Bayesian Nash equilibrium. Finally, given the profile of messages sent by the agents, the principal selects an agent according to the announced mechanism.

The paper proposes a mechanism which guarantees partial implementation for any knowledge network in which every agent is linked to at least one other agent. This means if every agent is linked to at least one other agent, then the proposed mechanism induces a static Bayesian game for which there exists an equilibrium strategy profile such that the principal selects the best agent with probability 1. Moreover, the mechanism achieves full implementation for knowledge networks which are complete or a star. This means if the knowledge network is complete or a star, then every equilibrium strategy profile is such that the principal selects the best agent with probability 1. However, full implementation is not achieved for every knowledge network; the circle network of four agents is an example for this.

The paper in particular relates to the literature on mechanism design for allocation, implementation and persuasion problems with one principal, multiple agents and (partially) verifiable private information. In our model private information is partially verifiable because of the limits to lying. An agent's message space varies with the true values about which she holds private information. Contributions from the mechanism design literature which are most closely related to this paper are the following. Lipman and Seppi (1995) analyze how much provability is needed to obtain "robust full revelation" when symmetrically informed agents with conflicting preferences communicate sequentially with an uninformed principal. Glazer and Rubinstein (2001) investigate which rules of debate maximize the probability that an uninformed principal chooses the correct one of two possible outcomes after listening to a debate between two fully informed agents. The two debaters have conflicting interests and can each raise a limited number of arguments. Deneckere and Severinov (2008) study partial implementation of social choice functions when each agent's message space depends on her type and is thus partial evidence of her type. Contrary to our paper, agents can only prove something about their own type, but not about other agents' types. In our setup, it is impossible to always identify the best agent without the *references* which are partial proofs about other agents' types. Ben-Porath and Lipman (2012) study full implementation of social choice functions with monetary transfers when agents can present evidence about the state. In allocation problems they fully implement certain social choice functions without monetary transfers but just by allocating a positive amount to every agent in equilibrium. Such a mechanism cannot "allocate" the prize to the best agent with probability 1. Kartik and Tercieux (2012) analyze which social choice functions are fully implementable when agents can send hard evidence and have full information about the state, in an otherwise general setting. Their results substantially rely on the full information of all agents, whereas a key aspect of our model is that agents can have different degrees of information as determined by their network position.

Within the literature on networks, Renou and Tomala (2012) study implementation of incentive-compatible social choice functions for different communication systems among agents and the mechanism designer. A network between agents and the designer determines the communication channels; an agent's ex-ante information is independent of her network position and her message space is determined by the designer. In our paper, an agent's information and message space are dependent on her network position and the communication system between agents and the designer is fixed; every agent can only communicate to the principal independent of the network architecture.

Several papers have investigated strategic communication and information transmission among agents in communication networks. Examples are Hagenbach and Koessler (2010), Garcia (2012), Galeotti et al. (2013), Acemoglu et al. (2014), Bloch et al. (2014), Patty and Penn (2014), Calvó-Armengol et al. (2015), Currarini and Feri (2015), Förster (2015), and Wolitzky (2015). These are different from the present paper, as we do not study strategic communication and information transmission among agents in a communication network, but between agents who have a knowledge network and an uninformed principal who wants to extract information from the network.

In section 2 and 3, we introduce the model and discuss the partial verifiability of information. In section 4, we construct the mechanism and provide examples to highlight the role of references and of the truth-telling assumption. In section 5, we first propose a strategy profile of agents, second, we show that given the strategy profile the constructed mechanism selects the best agent with certainty, and third, we prove that the proposed strategy profile is a Bayesian Nash equilibrium of the static game among agents. In section 6, it is shown that if the knowledge network is complete or a star, then every Bayesian Nash equilibrium of the game induced by the mechanism is such that the best agent is selected with certainty. However, we also provide an example of a knowledge network for which there exists an equilibrium such that the best agent is not selected with certainty. Finally, we conclude.

2 The Model

A principal has to assign a prize to one agent out of a set N of n agents where $n \geq 3$. Agents are competing for the prize because the prize is valuable to every agent. Each agent is differently suited to receive the prize. The principal knows which properties the ideal candidate should have. An agent i's value for the principal is measured by her absolute distance, d_i, to the ideal. Thus an agent matching the ideal would have $d_i = 0$. Regarding the distribution of distances across agents, we assume that each agent's distance is a random draw from distribution with strictly positive density f over the unit interval $[0, 1]$.[1] The profile of distances will be denoted by $d = (d_1, ..., d_n)$. The lower d_i, the better agent i fits the principal's purpose. It is further assumed that no two agents have the same distance. This means $d_i \neq d_j$ for $i \neq j$.[2] The principal does not know the distance of any agent but only the distribution of distances. Every agent exactly knows her own distance.

Agents are organized in a network characterized by the set of links L which exist among them. A link between agent i and agent j is denoted by ij. We say that two agents i and j are linked if link $ij \in L$. The network is undirected, thus if $ij \in L$, then $ji \in L$. An agent j to whom agent i is linked is called a *neighbor* of agent i. The set of all neighbors of agent i is $N_i := \{j \,|ij \in L\}$. Links signify familiarity: Besides knowing her own exact distance, an agent also knows the exact distances of her neighbors. Regarding the distances of agents who are not her neighbors, agent i only knows the distribution of distances across these agents. Importantly, we assume in the following that every agent has at least one neighbor. This is an important assumption as for identifying the best agent for all d it will be necessary that every agent is known to at least one other agent. Note that this assumption does not require the network to be connected. The network is common knowledge. This means everybody knows about which other agents j an agent i is informed.

For given d and L, agents do not only differ in their own distances but also in who their neighbors are and which distances their neighbors have. These aspects together determine an agent's type. Consequently, an agent i's type t_i is defined as the tuple of her own distance and all her neighbors' distances, $t_i := \left(d_i, d_{j_1}, ..., d_{j_{|N_i|}}\right)$ with $j_k \in N_i$ and $j_k \neq j_{k+1}$, and the type profile as $t := (t_1, ..., t_n)$.

[1] The assumption that distances have an upper bound is without loss of generality. All our results would still hold without assuming an upper bound.

[2] The assumption that no two agents have the same distance is for analytical convenience only. With this assumption we exclude special cases which would complicate the analysis significantly, and would not add further insights. Although we do not show it in this paper, we strongly believe that our results continue to hold if we allowed the distances of two agents to be the same.

After distances and thus types have realized, every agent sends a costless, private message about her information to the principal. The message from agent i is denoted by m_i and the profile of messages from all agents by $m := (m_1, ..., m_n)$. The profile of messages from all agents except i will be denoted by m_{-i}. Agent i's message contains statements about her own distance and about the distance of each of her neighbors. The statement about her own distance is m_{ii} and is called *application*; the statement about the distance of her neighbor $j \in N_i$ is m_{ij} and is called agent i's *reference* about j. Thus, $m_i = \left(m_{ii}, m_{ij_1}, ..., m_{ij_{|N_i|}}\right)$ with $j_k \in N_i$ and $j_k \neq j_{k+1}$. Agents can lie about their own distance and about a neighbor's distance but only to a certain extent: It is common knowledge that each statement m_{ik} must be an element of $[\max\{0, d_k - b\}, \min\{d_k + b, 1\}]$ where $b \in (0, \frac{1}{2})$ is the maximum possible lie about the distance of agent k. Hence, an agent's message space depends on her type.

The principal chooses a function π which specifies a probability π_i of being selected for every agent i for every possible m. More specifically, she chooses $\pi(m) = (\pi_1(m), ..., \pi_n(m))$ with $\pi_i(m) \in [0, 1]$ for every i and $\sum_{i \in N} \pi_i(m) = 1$ for every possible m. That fact that selection probabilities sum up to 1 given m accounts for the assumption that the principal has to assign the prize and cannot destroy it. The principal publicly announces π and commits to it before agents send their messages. The function π is the mechanism which the principal designs and commits to and according to which the principal selects an agent when she receives message profile m.

Knowing π, agents engage in a static Bayesian game. The solution concept applied will be Bayesian Nash equilibrium. Each agent chooses a message strategy $\hat{m}_i$ such that her message $\hat{m}_i(t_i)$ given any of her possible types t_i maximizes her expected probability of being selected, π_i^e, given every other agent j's message strategy, $\hat{m}_j$. The strategy profile will be denoted by $\hat{m}$ and the profile of every agent's strategy except agent i's by $\hat{m}_{-i}$. We introduce an assumption on an agent's preference for truth-telling. If the truthful message maximizes π_i^e for agent i of type t_i, then she tells the truth, this means $\hat{m}_{ik}(t_i) = d_k$ for all k. On the other hand, we will say that an agent with type t_i is lying if $\hat{m}_{ik}(t_i) \neq d_k$ for some k. Thus given her type an agent's first priority is to maximize her chances of being selected, and, second, she cares about truth-telling. After agents have chosen their strategies, distances and hence types realize. An agent of type t_i sends message $m_i = \hat{m}_i(t_i)$ and the message profile the principal receives is $m = \hat{m}(t)$.

The model summarizes as follows. First, the principal chooses, announces and commits to π. Second, agents engage in the static Bayesian game. They simultaneously choose their message strategies, then their types are realized and the according messages are privately send to the principal. Third, having received m the principal selects an agent according to $\pi(m)$.

The goal of this paper is to construct π such that the principal selects the best agent with probability 1. For any distance profile d, the *best* agent is agent i with $d_i = \min d$. The best agent will be called the *global minimum*.

In section 3, we briefly explain the partial verifiability of information in our model.

3 Partial Verifiability of Information

Before receiving the message profile m, the principal only knows the distribution of distances among agents. The message profile reveals partial information about the true distances of agents. The principal can infer from any message m_{ik} that $d_k \in [\max\{0, m_{ik} - b\}, \min\{m_{ik} + b, 1\}]$ because she is aware of the maximum possible lie b. This means that if $d_i < d_j$ for agent i and agent j and $ij \in L$, then i proves that she is better than j by sending $m_{ij} - m_{ii} > 2b$: The principal knows that $d_i \leq m_{ii} + b$ and $m_{ij} - b \leq d_j$. If $m_{ij} - m_{ii} > 2b \Leftrightarrow m_{ij} - b > m_{ii} + b$, then the principal also knows that $d_i \leq m_{ii} + b < m_{ij} - b \leq d_j$. Thus, in this case, the principal is sure that i is better than j. For certain configurations of d_i and d_j, i can always prove that she is better. For example, if $b < d_i < d_j < 1 - b$, i can always send m_{ii} and m_{ij} such that $m_{ij} - m_{ii} > 2b$.

Agent i does not prove that she is better than agent j, if $m_{ij} - m_{ii} < 2b$. In this case, the principal cannot be sure that i is better than j because $d_j < d_i$ with $d_i, d_j \in [\max\{0, m_{ij} - b\}, \min\{m_{ii} + b, 1\}]$ is possible. For certain configurations of d_i and d_j, i can never prove that she is better. For example, if $d_i < d_j < b$, i cannot send m_{ii} and m_{ij} such that $m_{ij} - m_{ii} > 2b$. She creates the largest difference between her application and her reference about j by choosing $m_{ii} = 0$ and $m_{ij} = d_j + b$ for which $m_{ij} - m_{ii} < 2b$ is true.

Observe that the partial verifiability of information due to the limits to lying implies that an agent i with type t_i can never imitate all other types $t_j \in [0, 1]$. Thus the revelation principle does not apply in this model.

In section 4, we first define the mechanism π. Second, we give examples to provide intuition for the mechanism and to highlight the role of references and of the truth-telling assumption.

4 The Mechanism

The following mechanism π defines the outcome of the static Bayesian game among agents for any message profile m with $\pi_i(m)$ being the probability for agent i to be selected given m. We show in section 5 that π as defined below induces a game for which there exists a Bayesian Nash equilibrium with message profile $m = \hat{m}(t)$ such that the principal selects the global minimum with probability 1 for any t. In section 6, we prove that every equilibrium of the game induced by π is such that the principal selects the global minimum with probability 1 for any t, if the knowledge network is complete or a star.

Definition 1 introduces π. The cardinality of a set x will be denoted by $|x|$.

Definition 1. *Define $\pi(m)$ for any message profile m as follows.*

When the principal receives message profile m, she first identifies all agents who send the best application. She defines these agents as set $B_1(m)$. Formally, $i \in B_1(m)$ if and only if $m_{ii} = \min_{j \in N} m_{jj}$.

Second, the principal selects a set $B_2(m)$ from $B_1(m)$.

If the best application is greater than zero, then all agents in $B_1(m)$ are in $B_2(m)$. Formally, if $\min_{j \in N} m_{jj} > 0$, then $B_2(m) = B_1(m)$.

If the best application is equal to zero, then the principal turns to the references about the agents in $B_1(m)$ to construct $B_2(m)$. For each $i \in B_1(m)$, the principal identifies the worst reference about i: $\overline{r}_i = \max_{j \in N_i} m_{ji}$. Then, the principal determines who receives the least bad worst reference (the min-max reference) among agents in $B_1(m)$. This means she determines every agent i for whom $\overline{r}_i = \min_{k \in B_1(m)} \overline{r}_k$. These agents then form the set $B_2(m)$. Formally, if $\min_{j \in N} m_{jj} = 0$, then $i \in B_2(m)$ if and only if $i \in B_1(m)$ and $\overline{r}_i = \min_{k \in B_1(m)} \overline{r}_k$.

Third, the principal selects a set $B_3(m)$ from $B_2(m)$. An agent $i \in B_2(m)$ is in $B_3(m)$ if and only if at least one of the following two conditions is satisfied:

1) Agent i's application conflicts with a reference about her or her reference about a neighbor conflicts with this neighbor's application, this means $m_{ii} \neq m_{ji}$ or $m_{ij} \neq m_{jj}$ for some $j \in N_i$.

2) Agent i's message proves that she is better than each of her neighbors, this means $m_{ij} - m_{ii} > 2b$ for all $j \in N_i$.

Having constructed $B_1(m)$, $B_2(m)$ and $B_3(m)$, the principal determines $\pi_i(m)$ for each $i \in N$.

If $B_3(m)$ is not empty, then the principal selects every $i \in B_3(m)$ with $\pi_i(m) = \frac{1}{|B_3(m)|}$.

If $B_3(m)$ is empty and $B_2(m)$ is not singleton, then the principal selects every $i \in B_2(m)$ with $\pi_i(m) = \frac{1}{|B_2(m)|}$.

If $B_3(m)$ is empty and $B_2(m) = \{i\}$, then the principal selects i with $\pi_i(m) = 0$. If there exists $j \notin N_i$, the principal selects every $j \notin N_i$ with $\pi_j(m) = \frac{1}{|N|-|N_i|-1}$. If every j is in N_i and there exists j with $m_{ij} - m_{ii} > 2b$, then the principal selects every j with this property with the same $\pi_j(m) = p$ such that $p|J| = 1$ where $J = \{j \in N_i | m_{ij} - m_{ii} > 2b\}$. If every j is in N_i and there does not exist $j \in N_i$ with $m_{ij} - m_{ii} > 2b$, then the principal selects every $j \in N_i$ with $\pi_j(m) = \frac{1}{|N|-1}$.

Observe that $B_2(m)$ is not empty for any m. There is always at least one agent who sends the best application among all agents such that $B_1(m)$ is not empty. Moreover, either $\min_i m_{ii} = 0$ or $\min_i m_{ii} > 0$ and at least one agent $i \in B_1(m)$ obtains $\overline{r}_i = \min_{k \in B_1(m)} \overline{r}_k$.

It is in order to provide some intuition for the mechanism and to highlight the importance of the truth-telling assumption. Assume the principal commits to π. Then, for any m, the principal first selects agents with the best application for $B_1(m)$; for $B_2(m)$, the principal selects all agents from $B_1(m)$, if the best application is larger than zero; if the best application is zero, then the principal compares the references about agents in $B_1(m)$ and selects those with the least bad reference for $B_2(m)$; for $B_3(m)$, the principal selects all agents from $B_2(m)$ who prove that they are better than all their neighbors, or who conflict with some neighbor. For every equilibrium which we establish in sections 5 and 6, it will always be the case that the global minimum is in $B_3(m)$. The different cases in which $B_3(m) = \emptyset$ serve as a punishment for agents who would deviate from equilibrium. For all of the following equilibria, $B_3(m) = \emptyset$ never occurs.

For the three examples below, assume $n = 3$ and the line network $L = \{12, 23\}$ as depicted in figure 1. In this network, 1 knows d_1 and d_2; 2 knows d_1, d_2 and d_3; 3 knows d_2 and d_3.

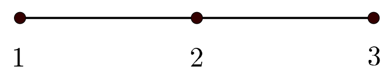

Figure 1: Line network with agents 1, 2 and 3.

Example 1. Consider $b = 0.2$ and the following realization of distances $d = (0.5, 0.6, 0.4)$ as depicted in figure 2 below.

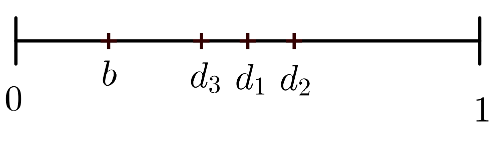

Figure 2: $b = 0.2$ and $d = (0.5, 0.6, 0.4)$ in the unit interval.

In this example, 3 is the global minimum. 1 and 3 both expect with positive probability to be the global minimum, and 2 knows that she is not the global minimum. Consider agents send the following message profile:

$m_{11} = d_1 - b = 0.3$, $m_{12} = d_2 + b = 0.8$;
$m_{22} = d_2 = 0.6$, $m_{21} = d_1 = 0.5$, $m_{23} = d_3 = 0.4$;
$m_{33} = d_3 - b = 0.2$, $m_{32} = d_2 + b = 0.8$.

1 and 3 exaggerate positively about themselves and negatively about 2, and 2 says the truth. The principal executes π such that $B_3(m) = \{3\}$ and $\pi_3(m) = 1$. No agent has an incentive to deviate given the other agents' strategies: 1 and 3 maximize their expected probability of being selected and truth-telling would be strictly worse for each of them in expectation. 2 knows that she has no chance of being selected and tells the truth. We will be precise about the corresponding equilibrium strategy profile and provide a proper proof in the next section. For now, this example should only serve to illustrate certain properties of the model and the mechanism. Observe that given m it is possible to differentiate between agents only on the basis of their applications. Applications are sufficient to identify the best agent in this case and $B_1(m) = B_2(m)$ is justified.

Example 2. Consider $b = 0.2$ and $d = (0.15, 0.3, 0.05)$ as in figure 3 below.

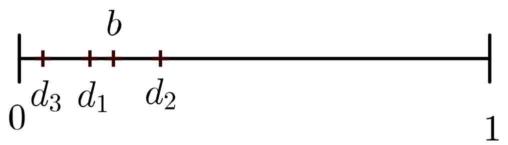

Figure 3: $b = 0.2$ and $d = (0.15, 0.3, 0.05)$ in the unit interval.

Let agents send the following message profile:

$m_{11} = 0$, $m_{12} = 0.5$;
$m_{22} = 0.3$, $m_{21} = 0.15$, $m_{23} = 0.05$;

$m_{33} = 0$, $m_{32} = 0.5$.

Again, 1 and 3 exaggerate positively about themselves and negatively about 2, and 2 says the truth. Executing π results into $B_3(m) = \{3\}$ and $\pi_3(m) = 1$. No agent has an incentive to deviate, for the same reasons as above. Observe that given m it is not possible to differentiate between 1 and 3 only on the basis of their application. Therefore the principal consults the references of 2 about 1 and 3. As 2 reports truthfully, the principal can now perfectly differentiate between 1 and 3. Thus, $B_1(m)$ still includes all agents with the best application which is zero in this case, but $B_2(m)$ only includes those who also get the least bad reference. In this case, it is 3.

Note that 2 knows that she is selected with probability zero for any of her messages, given the strategies of 1 and 3, because she cannot send an application of zero. Thus if we did not introduce the truth-telling assumption, 2 would be indifferent between any of her messages. The truth-telling assumption lets 2 strictly prefer the truth over any other message. The principal can rely on the informativeness of 2's references because 2 is not lying. If 2 was indifferent between all her messages, there would exist other equilibrium message profiles for which 2 would be lying. Then, the principal could not use 2's references to differentiate between 1 and 3 and she could not identify the best agent with certainty. Thus, if we dropped the truth-telling assumption, we would sacrifice the full implementation achieved in certain knowledge networks and could only sustain partial implementation.

The truth-telling assumption necessitates the introduction of $B_3(m)$. To see why, consider the next example.

Example 3. Consider $b = 0.2$ and $d = (0.5, 0.3, 0.4)$ as depicted in figure 4.

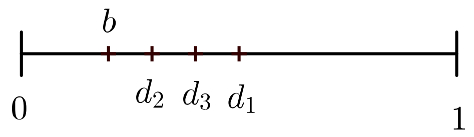

Figure 4: $b = 0.2$ and $d = (0.5, 0.3, 0.4)$ in the unit interval.

Thus 2 knows that she is the global minimum, and 1 and 3 know that they are not the global minimum. Let agents send the following message profile:

$m_{11} = 0.5$, $m_{12} = 0.3$;
$m_{22} = 0.1$, $m_{21} = 0.7$, $m_{23} = 0.6$;
$m_{33} = 0.4$, $m_{32} = 0.3$.

2 exaggerates positively about herself and negatively about 1 and 3, and 1 and 3 say the truth. Executing π results into $B_3(m) = \{2\}$ and $\pi_2(m) = 1$. No agent has an incentive to deviate, by the same arguments as above. To understand the role of $B_3(m)$, assume for a moment that $B_3(m)$ is not part of the mechanism, meaning that the principal ends the procedure after forming $B_2(m)$ and selects agents in $B_2(m)$ with equal probability. Then, for an agent to be selected, the principal only requires a best application, and, if this is zero, a least bad reference. In that case, given the above messages of 1 and 3, 2 is still selected with probability 1 if she deviates to the truth. Thus, given the truth-telling assumption, 2 has a strict incentive to deviate to the truth. If, however, all three agents tell the truth, 3 has an incentive to deviate and to send $m'_{33} = 0.2$.

The purpose of $B_3(m)$ is to force the global minimum to prove that she is better than all her neighbors or to conflict with some neighbor. We eliminate her incentive to deviate to the truth. Given the truth-telling assumption which is necessary for full implementation with our mechanism, $B_3(m)$ is needed to guarantee equilibrium existence.

In section 5, we propose a strategy profile of the game induced by π for n agents and for any network architecture. Subsequently, we show that the proposed strategy profile is a Bayesian Nash equilibrium of the game. The message profiles we used in the above examples correspond to this equilibrium strategy profile for $n = 3$ in the line network given each respective d.

5 Partial Implementation

In this section, we first define a strategy profile $\hat{m}$ for the static game following the announcement of π. Second, we show that if agents use $\hat{m}$ and the principal executes π, then the principal selects the best agent with certainty. Third, we show that $\hat{m}$ is an equilibrium given π. Remember that by assumption, the following analysis and results only apply to knowledge networks in which every agent has at least one neighbor, as mentioned in section 2.

For the definition of $\hat{m}$, we introduce three type categories according to which we can classify all possible types.

An agent i who is better than all of her neighbors and who is not linked to all other agents will be called a *local minimum with partial information*. Thus, these are all agents with a type t_i such that $|N_i| < n-1$ and $d_i < d_j$ for all $j \in N_i$. Such agents have *partial information* because they do not know every other agent. Agents 1 and 3 are both a local minimum with partial information in example 1 and 2.

An agent i who is better than all of her neighbors and who is linked to all other agents will be called a *local minimum with full information*. Thus, these are all agents with a type t_i such that $|N_i| = n - 1$ and $d_i < d_j$ for all $j \in N_i$. Such agents have *full information* because they know the distance of every other agent. Agent 2 is a local minimum with full information in example 3.

Note that the global minimum is a local minimum. A local minimum with partial information might be the global minimum, and a local minimum with full information is always the global minimum.

The third type category includes every agent who has at least one neighbor who is better than her. These are all agents with a type t_i for which $d_j < d_i$ for some $j \in N_i$. An agent in this category will be called *non-minimal.* Agent 2 is non-minimal in examples 1 and 2, and agents 1 and 3 are non-minimal in example 3.

In the following, an agent's message will only depend on her type category. Roughly speaking, we condense infinitely many types to just three types. At this point, we redefine an agent's type t_i as

$$t_i \in \{\textit{local minimum with part. inf.}, \textit{local minimum with full inf.}, \textit{non-minimal}\}\,.$$

Intuitively, the strategy profile $\hat{m}$ introduced below in definition 2 is such that every local minimum either proves that she is better than each of her neighbors or at least lies to the full extent about herself and her neighbors. Every non-minimal agent says the truth.

Definition 2. *The strategy profile $\hat{m}$ with $\hat{m}_i$ for every agent i is such that*

- *an agent who is non-minimal says the truth.*

  Formally, for $t_i =$ non-minimal

  $\hat{m}_{ii}(t_i) = d_i$ *and* $\hat{m}_{ij}(t_i) = d_j$ *for all* $j \in N_i$.

- *an agent i who is a local minimum with partial information sends the best possible application. Regarding her neighbors, she sends the worst possible references if some neighbor's distance is relatively close to her's, and she sends truthful references if all neighbors are sufficiently worse than her.*

  Formally, for $t_i =$ local minimum with partial information

  $\hat{m}_{ii}(t_i) = \max\{0, d_i - b\}$,

  $\hat{m}_{ij}(t_i) = \min\{d_j + b, 1\}$ *for all* $j \in N_i$*, if* $d_j - d_i \leq 2b$ *for some* $j \in N_i$,

  and $\hat{m}_{ij}(t_i) = d_j$ *for all* $j \in N_i$*, if* $d_j - d_i > 2b$ *for all* $j \in N_i$.

- *an agent who is a local minimum with full information sends the best possible application and the worst possible references if some neighbor's* *distance is sufficiently close to her's. If all her neighbors are sufficiently worse than her, she says the truth about herself and her neighbors.*

Formally, for $t_i =$ local minimum with full information

$\hat{m}_{ii}(t_i) = \max\{0, d_i - b\}$ *and* $\hat{m}_{ij}(t_i) = \min\{d_j + b, 1\}$ *for all* $j \in N_i$*, if* $d_j - d_i \leq 2b$ *for some* $j \in N_i$*,*

and $\hat{m}_{ii}(t_i) = d_i$ *and* $\hat{m}_{ij}(t_i) = d_j$ *for all* $j \in N_i$*, if* $d_j - d_i > 2b$ *for all* $j \in N_i$*.*

Proposition 5.1. *If agents use strategy profile* $\hat{m}$ *and the principal selects an agent according to* π*, then the principal selects the global minimum with probability 1 for every* t*.*

To prove that proposition 5.1 is true, we have to show that the claim is true for a) all t in which the global minimum is a local minimum with full information, and b) all t in which the global minimum is a local minimum with partial information. Case a) is reflected in example 3 in which agent 2 is the global minimum and is linked to all other agents. Case b) is reflected in examples 1 and 2 in which agent 3 is the global minimum and is not linked to all other agents. In the following proof, we first consider a) and second b).

Proof. Let agent g be the global minimum. If g is a local minimum with full information, then g is linked to every other agent $j \neq g$. Thus every other agent j is non-minimal, because she is linked to the better agent g. Then, every j sends the truthful application $m_{jj} = d_j$. Agent g either sends application $m_{gg} = \max\{0, d_g - b\}$ or $m_{gg} = d_g$. In each case, $m_{gg} < m_{jj}$ because $d_g < d_j$ for all j. Then $B_1(m) = \{g\}$.

As $B_2(m) \subseteq B_1(m) = \{g\}$ and $B_2(m)$ is not empty, it must be that $B_2(m) = B_1(m) = \{g\}$.

It is left to check if g is in $B_3(m)$. If all j are sufficiently worse than g, this means $d_j - d_g > 2b$ for all j, then g says the truth. The truthful message proves that she is better than each of her neighbors as $m_{gj} - m_{gg} > 2b$ for all j. Thus $B_3(m) = \{g\}$ and $\pi_g(m) = 1$. If some j's distance is sufficiently close to g's, this means $d_j - d_g \leq 2b$ for some j, then g sends the best possible application $m_{gg} = \max\{0, d_g - b\}$ and the worst possible reference $m_{gj} = \max\{d_j + b, 1\}$ for all j. Thus g's message conflicts, this means $m_{gg} \neq m_{jg}$ or $m_{gj} \neq m_{jj}$, with j for whom $d_j - d_g \leq 2b$. This ensures that $B_3(m) = \{g\}$ and $\pi_g(m) = 1$.

If g is a local minimum with partial information, then g is not linked to every other agent $j \neq g$ and every agent $j \neq g$ is either another local minimum with partial information or non-minimal.

Agent g and every other local minimum $j \neq g$ send their best possible applications $m_{gg} = \max\{0, d_g - b\}$ and $m_{jj} = \max\{0, d_j - b\}$, respectively. Every non-minimal agent j sends a truthful application $m_{jj} = d_j$. If $m_{gg} > 0$, then $m_{gg} < m_{jj}$ as $d_g < d_j$ for all j. Thus $B_1(m) = \{g\}$ and $B_2(m) = B_1(m)$. If $m_{gg} = 0$, then agent g is in $B_1(m)$. Moreover, each local minimum j with $d_j \leq b$ also sends $m_{jj} = 0$ and is in $B_1(m)$. As every agent in $B_1(m)$ is a local minimum, every neighbor k of an agent in $B_1(m)$ is non-minimal and says the truth. Thus all references about each $i \in B_1(m)$ are truthful and the worst reference about any $i \in B_1(m)$ is $\overline{r}_i = d_i$. As $d_g < d_j$ for all j, agent g receives the min-max reference of all agents in $B_1(m)$ and hence $B_2(m) = \{g\}$.

It is left to check if g is in $B_3(m)$. If $d_k - d_g > 2b$ for all $k \in N_g$, g sends $m_{gk} = d_k$. Then $m_{gk} - m_{gg} > 2b$ for all $k \in N_g$ and $B_3(m) = \{g\}$ and $\pi_g(m) = 1$. If $d_k - d_g \leq 2b$ for some $k \in N_g$, g sends $m_{gk} = \max\{d_k + b, 1\}$ for all $k \in N_g$. Then $m_{gg} \neq m_{jg}$ or $m_{gk} \neq m_{kk}$ for $k \in N_g$ for whom $d_k - d_g \leq 2b$. Hence, $B_3(m) = \{g\}$ and $\pi_g(m) = 1$. □

Next it is shown that the strategy profile $\hat{m}$ is a Bayesian Nash equilibrium of the static game following the announcement of π. First, we define the equilibrium conditions for our model.

Definition 3. *The strategy profile $\hat{m}$ is an equilibrium if for every agent i and for all t_i*

1. *$\hat{m}_i(t_i)$ maximizes π_i^e given $\hat{m}_{-i}$, and*

2. *$\hat{m}_i(t_i)$ is the truthful message if the truthful message maximizes π_i^e given $\hat{m}_{-i}$.*

If $\hat{m}_i$ is such that both of the above conditions which we will call 3.1. and 3.2., respectively, are satisfied for all t_i for every agent i, then no agent i with preferences as outlined in section 2 has an incentive to deviate from $\hat{m}_i$ given all other agents' strategies. Thus, if $\hat{m}_i$ satisfies 3.1. and 3.2. for all t_i for every agent i, then $\hat{m}$ is a Bayesian Nash equilibrium strategy profile.

Proposition 5.2. *The strategy profile $\hat{m}$ is an equilibrium strategy profile of the static Bayesian game induced by π.*

To prove proposition 5.2, we establish the following three lemmata each of which claims that, given one of the three possible types of agent i, $\hat{m}_i(t_i)$ maximizes π_i^e and is the truth if the truth maximizes π_i^e given $\hat{m}_{-i}$.

Lemma 5.3 deals with $t_i =$ *loc. min. with full inf.*, lemma 5.4 with $t_i =$ *loc. min. with partial inf.*, and lemma 5.5 with $t_i =$ *non-minimal.* These lemmata taken together prove that $\hat{m}$ is an equilibrium strategy profile.

Lemma 5.3. *If* $t_i =$ loc. min. with full inf., *then $\hat{m}_i(t_i) = m_i$ maximizes π_i^e and is the truth if the truth maximizes π_i^e given $\hat{m}_{-i}$.*

Proof. Let agent i be a local minimum with full information. This means agent i is linked to every other agent and each of her neighbors j is non-minimal with $d_j > d_i$. Thus, i knows that she is the global minimum and that $\pi_i(m) = 1$. There does not exist another message m_i' which increases π_i^e.

To prove the second part of the lemma we must show that if m_i is not the truth, then the truth does not maximize π_i^e. If agent g does not say the truth, then $d_j - d_i \leq 2b$ for some neighbor $j \in N_i$. Every $j \in N_i$ says the truth. If agent i deviates to the true message m_i', then $m_{ij}' = m_{jj}$ and $m_{ii}' = m_{ji}$ for all $j \in N_i$. Moreover, $m_{ij}' - m_{ii}' \leq 2b$ for $j \in N_i$ for whom $d_j - d_i \leq 2b$. Then $B_3(m_i', m_{-i}) = \emptyset$ and $B_2(m_i', m_{-i}) = \{i\}$ such that agent i is selected with probability 0. Thus, if m_i is not the truth and i deviates to the truth, then $\pi_i^e(m_i', m_{-i}) = 0 < \pi_i^e(m)$ and the truth does not maximize π_i^e.

□

Lemma 5.4. *If* $t_i =$ loc. min. with part. inf.*, then* $\hat{m}_i(t_i) = m_i$ *maximizes* π_i^e *and is the truth if the truth maximizes* π_i^e *given* $\hat{m}_{-i}$*.*

Proof. Let agent i be a local minimum with partial information. Then agent i expects with positive probability $p > 0$ to be the global minimum and that $\pi_i(m) = 1$. Thus, if she is the global minimum, m_i maximizes π_i.

With $1-p$, she expects not to be the global minimum and that $\pi_i(m) = 0$. If agent i is not the global minimum, then another agent $g \notin N_i$ is the global minimum with $d_g < d_i$. Agent g sends the best application $m_{gg} = \max\{0, d_g - b\} \leq m_{ii} = \max\{0, d_i - b\}$. The references about agent g and about agent i are given by non-minimal agents j only and are $m_{jg} = d_g$ and $m_{ji} = d_i$, respectively, for all j. Thus, for any m_i', agent g still sends the best application, receives the min-max reference among agents in $B_1(m_i', m_{-i})$ in case $m_{gg} = 0$, and conflicts with every neighbor $j \in N_g$ for whom $d_j - d_g \leq 2b$. Hence for any $m_i' \neq m_i$, $B_3(m_i', m_{-i}) = \{g\}$ and $\pi_i(m_i', m_{-i}) = 0$. Thus, if agent i is not the global minimum, m_i maximizes π_i.

As m_i maximizes π_i also if agent i is not the global minimum, m_i maximizes π_i^e for agent i.

Next, we show that π_i^e strictly decreases if m_i is not the truth and agent i deviates to the truth m_i'. Agent i expects with strictly positive probability that she is the global minimum, that $\pi_i(m) = 1$ and that there exists another local minimum j with $d_j \in (d_i, d_i + b)$ who sends $m_{jj} = \max\{0, d_j - b\}$. Consider i is indeed the global minimum and that such j exists. If $d_i > 0$ and i deviates from $m_{ii} = \max\{0, d_i - b\}$ to $m_{ii}' = d_i$, then her application is worse than j's. Thus $B_3(m_i', m_{-i}) \neq \emptyset$ but $i \notin B_3(m_i', m_{-i})$ such that $\pi_i(m_i', m_{-i}) = 0$. If $d_i = 0$, agent i only does not say the truth in case she has a neighbor k with $d_k - d_i \leq 2b$. By deviating to the truth, $B_3(m_i', m_{-i}) = \emptyset$ and $B_2(m_i', m_{-i}) = \{i\}$ such that $\pi_i(m_i', m_{-i}) = 0$.

□

Lemma 5.5. *If* $t_i =$ non-minimal*, then* $\hat{m}_i(t_i) = m_i$ *which is the truth maximizes* π_i^e*.*

Proof. Let agent i be non-minimal. Agent i knows that she is not the global minimum and that $\pi_i(m) = 0$ because she has a neighbor j with $d_j < d_i$.

We show that there does not exist a message m'_i for which $\pi^e_i(m'_i, m_{-i}) > 0$. This means that the true message m_i maximizes π^e_i.

Agent i is aware that the global minimum agent g sends application $m_{gg} = \max\{0, d_g - b\}$, unless g has full information and $d_j - d_g > 2b$ for all $j \in N_g$ in which case g sends $m_{gg} = d_g$. Remember that $i \in N_g$ if g has full information.

If $d_i > b$, then agent i's best feasible application is $d_i - b > 0$ which is strictly larger than $\max\{0, d_g - b\}$, or d_g if $d_i - d_g > 2b$. Hence, there is no m'_i such that $i \in B_1(m'_i, m_{-i})$.

If $d_i \leq b$, then agent i can choose $m'_{ii} = 0$ and, consequently, will be in $B_1(m'_i, m_{-i})$. However, there is no message m'_i such that i will be in $B_2(m'_i, m_{-i})$: The maximum reference about agent i is $\overline{r}_i = d_i + b$ if some neighbor of agent i is a local minimum and is $\overline{r}_i = d_i$ if no neighbor of agent i is a local minimum. Every neighbor of the global minimum is non-minimal, sends reference $m_{jg} = d_g$ about g and hence $\overline{r}_g = d_g$. If agent i is not a neighbor of g, then agent i cannot increase the maximum reference about g and $\overline{r}_g = d_g < \overline{r}_i$ for any message m'_i. If agent i is a neighbor of g, then agent i can increase the maximum reference about g to at most $m'_{ig} = d_g + b$. However, $\overline{r}_g = d_g + b$ is still less than $\overline{r}_i = d_i + b$ which is agent i's maximum reference as a neighbor of g. Thus there is no message m'_i such that $i \in B_2(m'_i, m_{-i})$ and gets selected with $\pi_i(m'_i, m_{-i}) > 0$ as a member of $B_2(m'_i, m_{-i})$ or $B_3(m'_i, m_{-i})$.

If i does not get selected with $\pi_i(m'_i, m_{-i}) > 0$ as a member of $B_2(m'_i, m_{-i})$ or $B_3(m'_i, m_{-i})$, the other possibility for i to get selected with $\pi_i(m'_i, m_{-i}) > 0$ is when $B_3(m'_i, m_{-i}) = \emptyset$ and $B_2(m'_i, m_{-i}) = \{j\}$ where $j \neq i$. Next, we show that there does not exist m'_i such that i gets selected with $\pi_i(m'_i, m_{-i}) > 0$ if $B_3(m'_i, m_{-i}) = \emptyset$ and $B_2(m'_i, m_{-i}) = \{j\}$ where $j \neq i$.

An agent j is in $B_2(m'_i, m_{-i})$ and not in $B_3(m'_i, m_{-i})$ if and only if, given (m'_i, m_{-i}), she sends the best application, receives the min-max reference among all agents in $B_1(m'_i, m_{-i})$ in case the best application is zero, and does not conflict with any neighbor where $m_{jk} - m_{jj} \leq 2b$ for some $k \in N_j$. Since agent i did sent $m_{ii} > \min_{k \in N} m_{kk}$ and cannot influence the application of other agents, the only candidate agents j to be in $B_2(m'_i, m_{-i})$ are those who already did sent $m_{jj} = \min_{k \in N} m_{kk}$ before any deviation of agent i.

These candidates are the global minimum g and other local minima l.

Before any deviation of i, g and every l conflict with each of their neighbors $k \in N_j$ with $j = l, g$ for whom $m_{jk} - m_{jj} \leq 2b$. In order for $j \in B_2(m_i', m_{-i})$ and $j \notin B_3(m_i', m_{-i})$, agent i must deviate to m_i' such that agent j does not conflict any more with any of her neighbors $k \in N_j$ and $m_{jk} - m_{jj} \leq 2b$ for some k. For such m_i' to exist, agent i must be a neighbor of agent j and $m_{ji} - m_{jj} \leq 2b$. Assume such m_i' exists.

If there exists some agent $k \neq i$ who is not a neighbor of j, then still $\pi_i(m_i', m_{-i}) = 0$.

If every agent $k \neq i$ is a neighbor of j and there is some k for whom $m_{jk} - m_{jj} > 2b$, then still $\pi_i(m_i', m_{-i}) = 0$.

If every agent $k \neq i$ is a neighbor of j and $m_{jk} - m_{jj} \leq 2b$ for all k, then such m_i' cannot exist: By assumption, $|N_j| \geq 2$ because $|N| \geq 3$. So even if agent i chooses a message such that she does not conflict with agent j, there is at least one other agent k with whom agent j is conflicting and j cannot be in $B_2(m_i', m_{-i})$ without being in $B_3(m_i', m_{-i})$.

Thus, there does not exist m_i' for which $B_3(m_i', m_{-i}) = \emptyset$, $B_2(m_i', m_{-i}) = \{j\}$ and $\pi_i(m_i', m_{-i}) > 0$.

Hence, $\pi_i^e(m_i', m_{-i}) = 0$ for every $m_i' \neq m_i$ and the true message m_i maximizes π_i^e. □

Lemma 5.3, 5.4, 5.5 together imply that $\hat{m}_i$ satisfies conditions 3.1. and 3.2. for all t_i for every agent i and thus $\hat{m}$ is a Bayesian Nash equilibrium strategy profile. Then proposition 5.2 is true.

6 Full Implementation

In this section, we first show that the complete network and the star network are two architectures of the knowledge network for which every equilibrium of the Bayesian game induced by π is such that the principal selects the global minimum with probability 1. Second, we show that the circle network of four agent is an example for a network architecture for which the principal does not select the global minimum with probability 1 in every equilibrium given π.

The knowledge network is complete if $ij \in L$ with $i \neq j$ for all $i \in N$ and all $j \in N$. The knowledge network is a star if there exists one agent $i \in N$ such that $ij \in L$ for all $j \in N$ and $jk \notin L$ for all $k \neq i$. We will call the agent who is linked to every other agent in the star network the center c. The knowledge network is a circle of four agents if it takes the following form:

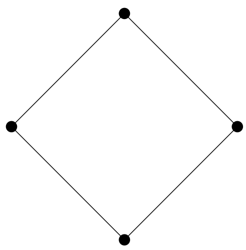

Figure 5: Circle network of four agents.

Proposition 6.1. *If the knowledge network is complete, then every equilibrium of the Bayesian game induced by π is such that the principal selects the global minimum with probability 1.*

Proof. Let the knowledge network be complete. As every agent has full information, the global minimum agent g knows t, m, and $\pi(m)$. Suppose that $\hat{m}$ is an equilibrium strategy profile and that $\hat{m}(t) = m$ and $\pi_g(m) < 1$ for t.

If $d_g > b$, then g can increase π_g by deviating to m'_g with $m'_{gg} = d_g - b$ and $m'_{gj} \neq m_{jj}$ for some $j \neq g$. After deviating, g uniquely sends the best application and conflicts with some neighbor. Hence, $B_3(m'_g, m_{-g}) = \{g\}$ and $\pi_g(m'_g, m_{-g}) = 1$.

If $d_g \leq b$, then g can increase π_g by deviating to m'_g with $m'_{gg} = 0$ and $m'_{gj} > d_g + b$ such that $m'_{gj} \neq m_{jj}$ for all $j \neq g$. After deviating, g sends the best application which is equal to zero and uniquely receives the min-max reference because g's reference about every agent $j \neq g$ is strictly greater than any reference about g. Moreover, g conflicts with her neighbors. Hence, $B_3(m'_g, m_{-g}) = \{g\}$ and $\pi_g(m'_g, m_{-g}) = 1$.

Then, however, $\hat{m}$ is not an equilibrium strategy profile because g can deviate to some m'_g which strictly increases π^e_g. $\square$

Proposition 6.2. *If the knowledge network is a star, then every equilibrium of the Bayesian game induced by π is such that the principal selects the global minimum with probability 1.*

Proof. Let the knowledge network be a star. First, we show by contradiction that there does not exist an equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $B_3(m) = \emptyset$. Second, we show by contradiction that every equilibrium strategy profile $\hat{m}$ is such that $m = \hat{m}(t)$ and $B_3(m) = \{g\}$ for all t. Given these two results, we then conclude that every equilibrium strategy profile $\hat{m}$ is such that $m = \hat{m}(t)$ and $\pi_g(m) = 1$ for all t.

First suppose that $\hat{m}$ is an equilibrium strategy profile and that $m = \hat{m}(t)$ and $c \in B_2(m)$ and $B_3(m) = \emptyset$ for t. From $c \in B_2(m)$ and $B_3(m) = \emptyset$ follows that $\pi_c(m) < 1$. Since c is linked to every other agent, she has full information and knows t, m and $\pi(m)$.

If $m_{cc} > \max\{d_c - b, 0\}$, then c strictly increases π_c by deviating to m'_c with $m'_{cc} < m_{cc}$ such that $m'_{cc} \neq m_{jc}$ for all $j \neq c$. After deviating, c uniquely sends the best application and conflicts with her neighbors. Hence, $B_3(m'_c, m_{-c}) = \{c\}$ and $\pi_c(m'_c, m_{-c}) = 1$.

If $m_{cc} = \max\{d_c - b, 0\}$ and $j \in B_2(m)$ for all $j \neq c$, then all $i \in N$ are in $B_2(m)$ and $B_3(m) = \emptyset$. Every agent sends the same best application and there are no conflicts, this means $m_{cc} = m_{jj} = m_{jc} = m_{cj}$ for all $j \neq c$. Then $\pi_c(m) = \frac{1}{n}$. In this case, c strictly increases π_c by deviating to m'_c with $m'_{cc} = m_{cc}$, $m'_{ck} \neq m_{ck}$ such that $m'_{ck} > 0$ for exactly one neighbor k and $m'_{cj} = m_{cj}$ for all other neighbors $j \neq k$. After c's deviation, all agents still send the same best application and there is exactly one conflict – the one between c and k. Thus, all $j \neq c, k$ are still in $B_2(m'_c, m_{-c})$ but not in $B_3(m'_c, m_{-c})$. If the best application is larger than zero, then $B_3(m'_c, m_{-c}) = \{c, k\}$. If the best application is zero and $\overline{r}_c = 0$ as $m_{jc} = 0$ for all $j \neq c$, then $B_3(m'_c, m_{-c}) = \{c\}$ because $\overline{r}_k = m'_{ck} > \overline{r}_c$. Thus $\pi_c(m'_c, m_{-c}) \geq \frac{1}{2} > \frac{1}{n}$.

If $m_{cc} = \max\{d_c - b, 0\}$ and some $j \notin B_2(m)$, then c strictly increases

π_c by deviating to m'_c with $m'_{cc} = m_{cc}$, $m'_{cj} \neq m_{cj}$ such that $m'_{cj} > 0$ for some $j \notin B_2(m)$, and $m'_{ck} = m_{ck}$ for all $k \in B_2(m)$. After c's deviation, still $B_2(m'_c, m_{-c}) = B_2(m)$ and now c is the only agent in $B_2(m'_c, m_{-c})$ who conflicts with a neighbor. Hence, $B_3(m'_c, m_{-c}) = \{c\}$ and $\pi_c(m'_c, m_{-c}) = 1$.

Then, however, $\hat{m}$ cannot be an equilibrium strategy profile because if $m = \hat{m}(t)$ and $c \in B_2(m)$ and $B_3(m) = \emptyset$ for t, then c can deviate to some m'_c which strictly increases π^e_c. Hence, there does not exist an equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $c \in B_2(m)$ and $B_3(m) = \emptyset$.

Second suppose that $\hat{m}$ is an equilibrium strategy profile and that $m = \hat{m}(t)$ and $j \in B_2(m)$ and $B_3(m) = \emptyset$ for t where j is not the center agent of the star. From $j \in B_2(m)$ and $B_3(m) = \emptyset$ follows that $\pi_j(m) \leq \frac{1}{2}$. Also $c \notin B_2(m)$ because $B_3(m) = \emptyset$, as we know from above, and $\pi_c(m) = 0$. As c has full information, c knows that $\pi_c(m) = 0$. From our equilibrium definition, it then follows that c must say the truth if $\hat{m}$ is an equilibrium profile. This implies that agent j also say the truth since she would otherwise conflict with c and the outcome would not be $j \in B_2(m)$ and $j \notin B_3(m) = \emptyset$.

Thus, for agent j who has partial information to expect with positive probability that $j \in B_2(m)$ and $B_3(m) = \emptyset$, her own message must be such that $m_{jj} = d_j$ and $m_{jc} = d_c$. We show that if $m_{jj} = d_j$ and $m_{jc} = d_c$ and j expects $j \in B_2(m)$ and $B_3(m) = \emptyset$ with positive probability, then j strictly increases π^e_j by deviating to another message m'_j. From this, we can conclude that $\hat{m}$ cannot be an equilibrium.

Let $m_{jj} = d_j$ and $m_{jc} = d_c$ and let j expect $j \in B_2(m)$ and $B_3(m) = \emptyset$ with positive probability.

Assume first that $d_j > 0$.

If it is indeed the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j strictly increases π_j by deviating to m'_j with $m'_{jj} < d_j$ such that $m'_{jj} \neq \hat{m}_{cj}(t_c)$ for all t_c and $m'_{jc} = d_c$. The deviation results into $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$.

If it is not the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j does not do worse by deviating to m'_j. If $j \in B_3(m)$, then $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$. If $B_2(m) = \{k\}$ where $k \neq j, c$ and $B_3(m) = \emptyset$, then either

$B_2(m'_j, m_{-j}) = B_2(m) = \{k\}$ and $\pi_j(m'_j, m_{-j}) = \pi_j(m)$ or $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$. In all other remaining sub-cases which could occur in equilibrium $\pi_j(m) = 0$ and j always weakly improves by deviating to m'_j. Remember, also for the following paragraphs, that we showed before that the sub-case of $B_2(m) = \{c\}$ and $B_3(m) = \emptyset$ which could imply $\pi_j(m) > 0$ does not occur in equilibrium.

Assume second that $d_j = 0$ and $d_c < 1$.

If it is indeed the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j strictly increases π_j by deviating to m'_j with $m'_{jj} = d_j$ and $m'_{jc} > d_c$ such that $m'_{jc} \neq \hat{m}_{cc}(t_c)$ for all t_c. The deviation results into $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$.

If it is not the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j does not do worse by deviating to m'_j. If $j \in B_3(m)$, then $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ and $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$. If $B_2(m) = \{k\}$ where $k \neq j, c$ and $B_3(m) = \emptyset$, then $B_2(m'_j, m_{-j}) = B_2(m)$ and $\pi_j(m'_j, m_{-j}) = \pi_j(m)$. For all other remaining sub-cases which could occur in equilibrium $\pi_j(m) = 0$ and thus j always weakly improves by deviating to m'_j.

Assume third that $d_j = 0$ and $d_c = 1$.

If it is indeed the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j strictly increases π_j by deviating to m'_j with $m'_{jj} = d_j$ and $m'_{jc} \in (2b, 1) < d_c$ such that $m'_{jc} \neq \hat{m}_{cc}(t_c)$ for all t_c. The deviation results into $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$.

If it is not the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then j does not do worse by deviating to m'_j. If $j \in B_3(m)$, then $B_3(m'_j, m_{-j}) = B_3(m)$ and $\pi_j(m'_j, m_{-j}) = \pi_j(m)$. If $B_2(m) = \{k\}$ where $k \neq j, c$ and $B_3(m) = \emptyset$, then $B_2(m'_j, m_{-j}) = B_2(m)$ and $\pi_j(m'_j, m_{-j}) = \pi_j(m)$. For all other remaining sub-cases which could occur in equilibrium $\pi_j(m) = 0$ and thus j always weakly improves by deviating to m'_j.

Then, however, $\hat{m}$ is not an equilibrium strategy profile because there always exists a deviation for j which strictly increases π^e_j if she expects with positive probability that $j \in B_2(m)$ and $j \notin B_3(m)$. Hence, there does not exist an equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $j \in B_2(m)$ and $B_3(m) = \emptyset$ where j is not the center of the star.

Thus, there is no equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $B_3(m) = \emptyset$. This completes the first part of the proof for proposition 6.2.

The implication of this first result is that for every equilibrium strategy profile $\hat{m}$ and every t such that $m = \hat{m}(t)$ it is true that $\pi_i(m) > 0$ if and only if $i \in B_3(m)$ for all $i \in N$.

Next, we show by contradiction that any equilibrium strategy profile $\hat{m}$ is such that $m = \hat{m}(t)$ and $B_3(m) = \{g\}$ for all t where g is the global minimum.

First, assume that $\hat{m}$ is an equilibrium strategy profile and that $m = \hat{m}(t)$ and $B_3(m) \neq \{c\}$ for t where c is the global minimum. From $B_3(m) \neq \{c\}$ follows that $\pi_c(m) < 1$. Since c has full information, she knows t, m and $\pi(m)$.

If $d_c > b$, then c can strictly increase π_c by deviating to m'_c with $m'_{cc} = d_c - b$ and $m'_{cj} \neq m_{jj}$ for all j. The deviation leads to $B_3(m'_c, m_{-c}) = \{c\}$ and $\pi_c(m'_c, m_{-c}) = 1$. If $d_c \leq b$, then c can strictly increase π_c by deviating to m'_c with $m'_{cc} = 0$ and $m'_{cj} > d_c + b$ such that $m'_{cj} \neq m_{jj}$ for all j. The deviation leads to $B_3(m'_c, m_{-c}) = \{c\}$ and $\pi_c(m'_c, m_{-c}) = 1$.

Then, however, $\hat{m}$ is not an equilibrium strategy profile. Hence, there does not exist an equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $B_3(m) \neq \{c\}$ where c is the global minimum.

Second, assume that $\hat{m}$ is an equilibrium strategy profile and that $m = \hat{m}(t)$ and $B_3(m) \neq \{j\}$ for some t where $j \neq c$ is the global minimum. From $B_3(m) \neq \{j\}$ follows that $\pi_j(m) < 1$. Given $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum, j is a local minimum with partial information and expects with positive probability that $B_3(m) \neq \{j\}$ and that she is the global minimum.

Let $j \neq c$ expect with positive probability to be the global minimum and that $B_3(m) \neq \{j\}$.

Assume first that $d_j > b$. If it is indeed the case that $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum, then j strictly increases π_j by deviating to m'_j with $m'_{jj} = d_j - b$ and $m'_{jc} \neq \hat{m}_{cc}(t_c)$ for all t_c. The deviation leads

$B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$. If it is not the case that $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum, then j does never do worse by deviating to m'_j. If $j \in B_3(m)$, then j still sends the best application which is greater than zero after her deviation to m'_j. Thus, $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ and $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$. If $j \notin B_3(m)$, then $\pi_j(m) = 0$. In this case deviating to m'_j always weakly increases π_j.

Assume second that $d_j \leq b < d_c$. Define agent j's deviation as m'_j with $m'_{jj} = 0$ and $m'_{jc} > d_j + b$ such that $m'_{jc} \neq \hat{m}_{cc}(t_c)$ for all t_c.

If $B_3(m) \neq \{j\}$ and $c \in B_3(m)$ where $j \neq c$ is the global minimum, then $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$ because j uniquely sends the best application after deviating.

If $B_3(m) \neq \{j\}$ and $c \notin B_3(m)$ where $j \neq c$ is the global minimum, then $\pi_c(m) = 0$. This implies that m_c must be the true message if $\hat{m}$ is an equilibrium profile because c has full information. If m_c is the truth and j deviates to m'_j, then $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$ because j as the global minimum receives the min-max reference from c.

Thus, if $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum, then j strictly improves π_j by deviating to m'_j.

If it is not the case that $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum, then j does never do worse by deviating to m'_j. If $j \in B_3(m)$, then $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ and $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$. If $j \notin B_3(m)$, then $\pi_j(m) = 0$ and j always weakly improves by deviating to m'_j.

Up to here, we showed that there does not exist an equilibrium strategy profile $\hat{m}$ and t such that $m = \hat{m}(t)$ and $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum and $d_j > b$ or $d_j \leq b < d_c$. The proof strategy was to define a profitable deviation for agent j, if $d_j > b$ or $d_j \leq b < d_c$ and she expected with positive probability to be the global minimum and that $B_3(m) \neq \{j\}$.

We use a different proof strategy to show that there does not exist an equilibrium $\hat{m}$ and t such that $m = \hat{m}(t)$ and $B_3(m) \neq \{j\}$ where $j \neq c$ is the global minimum and $d_j < d_c \leq b$. First, we show that every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ where an agent $j \neq c$ with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} > d_j + b$. From this, we derive that every equilibrium $\hat{m}$

and t is such that $m = \hat{m}(t)$ where c says the truth if there exists an agent j with $d_j < d_c \leq b$. If an agent j with $d_j < d_c \leq b$ exists, then such an agent j is the global minimum. From these three observations, we conclude that every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ and $B_3(m) = \{j\}$ where the global minimum is agent $j \neq c$ with $d_j < d_c \leq b$.

First, suppose $\hat{m}$ is an equilibrium and that $m = \hat{m}(t)$ for t such that agent j with $d_j < d_c \leq b$ sends $m_{jj} > 0$. Given $d_j < d_c \leq b$, agent j expects with strictly positive probability that $d_j < d_c \leq b < d_k$ for all $k \neq j$. If indeed $d_j < d_c \leq b < d_k$ for all $k \neq j$, then c sends m_c such that $B_3(m) = \{c\}$ in order to maximize π_c^e given $\hat{m}_{-c}$. Then, however, π_j strictly increases, if j deviates to m'_j with $m'_{jj} = 0$ and $m'_{jc} > d_j + b$ with $m'_{jc} \neq m'_{cc}(t_c)$ for all t_c such that $B_3(m'_j, m_{-j}) = \{j\}$. If it is not the case that $d_j < d_c \leq b < d_k$ for all $k \neq j$ and if $j \in B_3(m)$ which is necessary for $\pi_j(m) > 0$, then deviating to m'_j leads to $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$ as well. Thus, deviating to m'_j strictly increases π_j^e such that $\hat{m}$ is not an equilibrium. Hence, there is no equilibrium $\hat{m}$ and t such that $m = \hat{m}(t)$ where agent j with $d_j < d_c \leq b$ sends $m_{jj} > 0$.

Second, suppose $\hat{m}$ is an equilibrium and that $m = \hat{m}(t)$ for t such that agent j with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} < d_j + b$. Given $d_j < d_c \leq b$, agent j expects with strictly positive probability that $d_j < d_c \leq b < d_k$ for all $k \neq j$. If indeed $d_j < d_c \leq b < d_k$ for all $k \neq j$, then every agent k says the truth. If agent k did not say the truth, this means that she expected with positive probability to be $B_3(m)$ and that $\pi_k(m) > 0$. Suppose $k \in B_3(m)$ and that $\pi_k(m) > 0$ for agent k with $d_c \leq b < d_k$. Then c would always deviate to a message m'_c such that $B_3(m'_c, m_{-c}) = \{c\}$. Thus, agent k with $d_k > b \geq d_c$ expects with probability 0 to be in $B_3(m)$ in equilibrium and says the truth. Then, given m_{-c}, c sends $m_{cc} = 0$ and some $m_{cj} > m_{jc}$ such that $B_3(m) = \{c\}$ to maximize π_c^e. In this case, however, π_j strictly increases, if j deviates to m'_j with $m'_{jj} = 0$ and $m'_{jc} > d_j + b$ with $m'_{jc} \neq m'_{cc}(t_c)$ for all t_c such that $B_3(m'_j, m_{-j}) = \{j\}$. If it is not the case that $d_j < d_c \leq b < d_k$ for all $k \neq j$ and if $j \in B_3(m)$ which is necessary for $\pi_j(m) > 0$, then deviating to m'_j implies $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ and $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$ and weakly improves j as well. Thus, deviating to m'_j strictly increases π_j^e such that

$\hat{m}$ is not an equilibrium. Hence, there is no equilibrium $\hat{m}$ and t such that $m = \hat{m}(t)$ where agent j with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} < d_j + b$.

Third, suppose $\hat{m}$ is an equilibrium and that $m = \hat{m}(t)$ for t such that agent j with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} = d_j + b$. Given $d_j < d_c \leq b$, agent j expects with strictly positive probability that $d_j < d_c \leq b < d_k$ for all $k \neq j$. If indeed $d_j < d_c \leq b < d_k$ for all $k \neq j$, then every agent k says the truth. Thus, given m_{-c}, c sends $m_{cc} = 0$ and $m_{cj} = m_{jc}$ such that $B_3(m) = \{j, c\}$ to maximize π_c^e. Then, however, π_j strictly increases, if j deviates to m'_j with $m'_{jj} = 0$ and $m'_{jc} > d_j + b$ with $m'_{jc} \neq m'_{cc}(t_c)$ for all t_c such that $B_3(m'_j, m_{-j}) = \{j\}$. If it is not the case that $d_j < d_c \leq b < d_k$ for all $k \neq j$ and if $j \in B_3(m)$ which is necessary for $\pi_j(m) > 0$, then deviating to m'_j implies $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ and $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$ and weakly improves j as well. Thus, deviating to m'_j strictly increases π_j^e such that $\hat{m}$ is not an equilibrium. Hence, there is no equilibrium $\hat{m}$ and t such that $m = \hat{m}(t)$ where agent j with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} = d_j + b$.

Thus, every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ where an agent $j \neq c$ with $d_j < d_c \leq b$ sends $m_{jj} = 0$ and $m_{jc} > d_j + b$. This implies that every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ and $\pi_c(m) = 0$ if there exists an agent j with $d_j < d_c \leq b$ because the best application is zero and the worst reference about c is worse than any feasible reference about j. As c has full information and knows that $\pi_c(m) = 0$, she says the truth. Thus, every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ where c says the truth, if there exists an agent j with $d_j < d_c \leq b$.

If there exists an agent j with $d_j < d_c \leq b$, then such an agent j is the global minimum g. Thus, for every equilibrium $\hat{m}$ and t such that $m = \hat{m}(t)$ and such that there exists an agent j with $d_j < d_c \leq b$, the global minimum sends $m_{gg} = 0$ and receives the min-max reference since c says the truth. Moreover, the global minimum conflicts with c. Thus, $B_3(m) = \{g\}$.

Hence, every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$ and $B_3(m) = \{j\}$ where $j \neq c$ with $d_j < d_c \leq b$ is the global minimum.

Thus, if the knowledge network is a star, then every equilibrium $\hat{m}$ and t is such that $m = \hat{m}(t)$, $B_3(m) = \{g\}$ and $\pi_g(m) = 1$ where g is the global minimum. $\square$

The mechanism π does not guarantee full implementation for every knowledge network. The following example identifies a network architecture and an equilibrium $\hat{m}$ such that the principal does not select the global minimum with probability 1.

Example 4. Let $n = 4$ and let the network architecture be a circle as shown in figure 6.

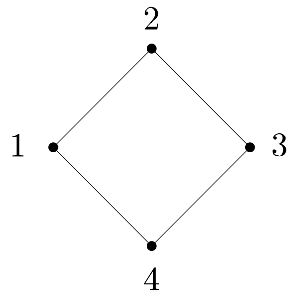

Figure 6: Circle network with $n = 4$.

Let $\hat{m}$ be such that

- a local minimum sends the best possible application. Regarding her neighbors, she sends the worst possible references if some neighbor's distance is relatively close to her's, and she sends truthful references if both neighbors are sufficiently worse than her.

  Formally, if $d_i < d_j$ for all $j \in N_i$, then

  $\hat{m}_{ii}(t_i) = \max\{0, d_i - b\}$,

  $\hat{m}_{ij}(t_i) = \min\{d_j + b, 1\}$ for all $j \in N_i$, if $d_j - d_i \leq 2b$ for some $j \in N_i$,

  and $\hat{m}_{ij}(t_i) = d_j$ for all $j \in N_i$, if $d_j - d_i > 2b$ for all $j \in N_i$.

- a non-minimal agent i who has $d_i > b$ or a neighbor j with $d_j = 0$ says the truth.

  Formally, if $d_i > d_j$ for some $j \in N_i$ and if $d_i > b$ or $d_j = 0$ for some $j \in N_i$, then

  $\hat{m}_{ii}(t_i) = d_i$ and $\hat{m}_{ij}(t_i) = d_j$ for all $j \in N_i$.

- a non-minimal agent i who has $d_i \leq b$ and no neighbor j with $d_j = 0$ sends the best possible application, a reference equal to b for every neighbor j with $d_j \leq b$ and a truthful reference about a neighbor with $d_j > b$.

  Formally, if $d_i > d_j$ for some $j \in N_i$ and if $d_i \leq b$ and $d_j = 0$ for no $j \in N_i$, then

  $\hat{m}_{ii}(t_i) = 0$,

  $\hat{m}_{ij}(t_i) = b$ for all $j \in N_i$ with $d_j \leq b$,

  and $\hat{m}_{ij}(t_i) = d_j$ for $j \in N_i$ with $d_j > b$.

Lemma 6.3. *If t_i is such that agent i is a local minimum, then $\hat{m}_i(t_i) = m_i$ maximizes π_i^e and is the truth if the truth maximizes π_i^e given $\hat{m}_{-i}$.*

Proof. First, we show that $\hat{m}_i(t_i) = m_i$ maximizes π_i^e given $\hat{m}_{-i}$, if t_i is such that agent i is a local minimum.

Let t_i be such that that agent i is a local minimum and let $m_i = \hat{m}_i(t_i)$.

Suppose $i = g$ and $j \notin N_i$ has $d_j > b$. Then $B_3(m) = \{i\}$ because $m_{ii} = \min_{k \in N} m_{kk}$, $\overline{r}_i = \min_{k \in N} \overline{r}_k$ and i conflicts with neighbor $j \in N_i$ if $d_j - d_i \leq 2b$. This implies $\pi_i(m) = 1$.

Suppose $i = g$ and $j \notin N_i$ has $d_j \leq b$. Then i and $j \notin N_i$ both send the best application 0. No $j \in N_i$ is in $B_2(m)$ because i sends reference $m_{ij} = \min\{d_j + b, 1\}$ about each $j \in N_i$ if one of them is sufficiently close to i. If both neighbors $j \in N_i$ of i send truthful references, then $B_3(m) = \{i\}$ and hence $\pi_i(m) = 1$. If one neighbor $j \in N_i$ sends references equal to b, then $B_3(m) = \{i, j\}$ for $j \notin N_i$ and $\pi_i(m) = \frac{1}{2}$. In this case there does not exist $m'_i \neq m_i$ such that $B_3(m'_i, m_{-i}) = \{i\}$ or $B_2(m'_i, m_{-i}) = \{j\}$ with $j \notin N_i$ and hence there exists no $m'_i \neq m_i$ such that $\pi_i(m'_i, m_{-i}) > \pi_i(m)$.

Suppose $i \neq g$. Then $j \notin N_i$ is g.

If $d_i > b$, then $B_3(m) = \{g\}$ and $\pi_i(m) = 0$. In this case there does not exist $m'_i \neq m_i$ such that $i \in B_1(m'_i, m_{-i})$ or $B_2(m'_i, m_{-i}) = \{j\}$ with $j \notin N_i$ and hence there exists no $m'_i \neq m_i$ such that $\pi_i(m'_i, m_{-i}) > \pi_i(m)$.

If $d_i \leq b$, and both neighbors $j \in N_i$ of i send truthful references, then $B_3(m) = \{g\}$ and $\pi_i(m) = 0$. In this case there does not exist $m'_i \neq m_i$ such that $i \in B_2(m_i', m_{-i})$ or $B_2(m_i', m_{-i}) = \{j\}$ with $j \notin N_i$ and hence there exists no $m_i' \neq m_i$ such that $\pi_i(m_i', m_{-i}) > \pi_i(m)$.

If $d_i \leq b$, and one neighbor $j \in N_i$ of i sends references equal to b, then $B_3(m) = \{g, i\}$ and $\pi_i(m) = \frac{1}{2}$. In this case there does not exist $m_i' \neq m_i$ such that $B_3(m_i', m_{-i}) = \{i\}$ or $B_2(m_i', m_{-i}) = \{j\}$ with $j \notin N_i$ and hence there exists no $m_i' \neq m_i$ such that $\pi_i(m_i', m_{-i}) > \pi_i(m)$.

Thus, $\hat{m}_i(t_i) = m_i$ maximizes π_i^e given $\hat{m}_{-i}$, if t_i is such that agent i is a local minimum.

Second, we show that if t_i is such that agent i is a local minimum and $\hat{m}_i(t_i) = m_i$ is not the truth, then deviating to the truth strictly decreases π_i^e given $\hat{m}_{-i}$.

A local minimum i with $d_i > 0$ assigns a strictly positive probability to $i = g$ and $j \notin N_i$ being a local minimum as well with $d_j \in (d_i, d_i + b)$. If indeed $i = g$ and $j \notin N_i$ is a local minimum as well with $d_j \in (d_i, d_i + b)$, then $i \in B_3(m)$ and $\pi_i(m) > 0$. In this case, a deviation by i to the truth results into $i \notin B_1(m_i', m_{-i})$ and $B_3(m_i', m_{-i}) = \{j\}$ with $j \notin N_i$ and thus $\pi_i(m_i', m_{-i}) = 0$.

A local minimum i with $d_i = 0$ only does not say the truth if $d_j - d_i \leq 2b$ for some $j \in N_i$. Given m, $B_3(m) = \{i\}$ and $\pi_i(m) = 1$. If i deviates to the truth, then $B_2(m_i', m_{-i}) = \{i\}$ and $B_3(m_i', m_{-i}) = \emptyset$ and thus $\pi_i(m_i', m_{-i}) = 0$.

Thus, if t_i is such that agent i is a local minimum and $\hat{m}_i(t_i) = m_i$ is not the truth, then deviating to the truth strictly decreases π_i^e given $\hat{m}_{-i}$.

$\square$

Lemma 6.4. *If t_i is such that agent i is non-minimal and has $d_i > b$ or a neighbor $j \in N_i$ with $d_j = 0$, then $\hat{m}_i(t_i) = m_i$ which is the truth maximizes π_i^e given $\hat{m}_{-i}$.*

Proof. Let t_i be such that agent i is non-minimal and has $d_i > b$ or a neighbor $j \in N_i$ with $d_j = 0$ and let $m_i = \hat{m}_i(t_i)$. Given m, $i \notin B_3(m) \neq \emptyset$ and thus $\pi_i^e(m) = 0$.

If $d_i > b$, then there is no $m_i' \neq m_i$ such that $i \in B_1(m_i', m_{-i})$ because $m_{gg} < d_i - b$. Moreover, there does not exist $m_i' \neq m_i$ such that $B_2(m_i', m_{-i}) = \{j\}$ with $j \notin N_i$ and $B_3(m_i', m_{-i}) = \emptyset$: If $j = g$ for $j \notin N_i$, then $B_3(m_i', m_{-i}) = \{g\}$ for all $m_i' \neq m_i$. If $j \neq g$ for $j \in N_i$, then $j \notin N_i$ is never in $B_2(m_i', m_{-i})$ for any $m_i' \neq m_i$ because g sends $m_{gg} = \max\{0, d_g - b\}$ and $m_{gj} = \min\{d_j + b, 1\}$ if $d_j - d_g \leq 2b$ for $j \in N_g$.

If i has a neighbor with $d_j = 0$, then this neighbor is g. There does not exist $m_i' \neq m_i$ such that $i \in B_2(m_i', m_{-i})$ because $m_{gg} = 0$ and $m_{gi} = \min\{d_i + b, 1\}$ if $d_i \leq 2b$. Moreover, $j \notin N_i$ is never in $B_2(m_i', m_{-i})$ for any $m_i' \neq m_i$ because g sends $m_{gg} = \max\{0, d_g - b\}$ and $m_{gj} = \min\{d_j + b, 1\}$ if $d_j - d_i \leq 2b$ for $j \in N_g$.

Thus, if t_i is such that agent i is non-minimal and has $d_i > b$ or a neighbor $j \in N_i$ with $d_j = 0$, then $\hat{m}_i(t_i) = m_i$ which is the truth maximizes π_i^e given $\hat{m}_{-i}$.

□

Lemma 6.5. *If t_i is such that agent i is non-minimal and has $d_i \leq b$ and no neighbor j with $d_j = 0$, then $\hat{m}_i(t_i) = m_i$ which is not the truth maximizes π_i^e and the truth does not maximize π_i^e given $\hat{m}_{-i}$.*

Proof. First, we show that $\hat{m}_i(t_i) = m_i$ which is not the truth maximizes π_i^e given $\hat{m}_{-i}$, if t_i is such that agent i is non-minimal, has $d_i \leq b$ and has no neighbor j with $d_j = 0$.

Let t_i be such that agent i is non-minimal, has $d_i \leq b$ and has no neighbor j with $d_j = 0$ and let $m_i = \hat{m}_i(t_i)$.

If $g \notin N_i$ and $d_g > 0$, then at least one neighbor of i and g sends references equal to b such that $B_3(m) = \{g, i\}$ and $\pi_i(m) = \frac{1}{2}$. For any $m_i' \neq m_i$, $g \in B_3(m_i', m_{-i})$ and thus no $m_i' \neq m_i$ increases π_i.

If $g \notin N_i$ and $d_g = 0$, then $B_3(m) = \{g\}$ and $\pi_i(m) = 0$. For any $m_i' \neq m_i$, $B_3(m_i', m_{-i}) = \{g\}$ because $m_{gg} = 0$ and the neighbors of i and g say the truth.

If $g \in N_i$ and $d_g > 0$, then $i \notin B_3(m) \neq \emptyset$ and $\pi_i(m) = 0$. There does not exist $m_i' \neq m_i$ such that $i \in B_2(m_i', m_{-i})$ because $m_{gg} = 0$ and $m_{gi} = d_i + b$. Moreover, there is no $m_i' \neq m_i$ such that $j \in B_2(m_i', m_{-i})$ for $j \notin N_i$ because

$m_{gg} = 0$ and $m_{gj} = \min\{d_j + b, 1\}$.

Thus, $\hat{m}_i(t_i) = m_i$ which is not the truth maximizes π_i^e given $\hat{m}_{-i}$, if t_i is such that agent i is non-minimal, has $d_i \leq b$ and has no neighbor j with $d_j = 0$.

Second, we show that if t_i is such that agent i is non-minimal, has $d_i \leq b$, has no neighbor j with $d_j = 0$ and $\hat{m}_i(t_i) = m_i$ which is not the truth, then deviating to the truth strictly decreases π_i^e given $\hat{m}_{-i}$.

Let t_i be such that agent i is non-minimal, has $d_i \leq b$ and has no neighbor j with $d_j = 0$ and let $m_i = \hat{m}_i(t_i)$.

Agent i assigns strictly positive probability to $g \notin N_i$ and $d_g > 0$. If $g \notin N_i$ and $d_g > 0$, then $B_3(m) = \{g, i\}$ and $\pi_i(m) = \frac{1}{2}$. Deviating to the true message results into $B_3(m_i', m_{-i}) = \{g\}$ and $\pi_i(m_i', m_{-i}) = 0$ because $m_{ii}' > m_{gg}$.

Thus, deviating to the truth strictly decreases π_i^e. $\square$

Lemmata 6.3, 6.4, and 6.5 together imply that $\hat{m}$ as defined in example 4 is an equilibrium strategy profile for the circle network with $n = 4$.

Consider the type realization in the circle network with $n = 4$ is such that all $i \in N$ have $d_i \leq b$ and $d_g > 0$ and that agents use the equilibrium $\hat{m}$ as defined in example 4. Then $B_3(m) = \{g, j\}$ where $j \notin N_g$ and $\pi_g(m) = \frac{1}{2}$. This case occurs with a strictly positive probability. Thus, $\hat{m}$ is an equilibrium strategy profile for the circle network with $n = 4$ for which the principal does not expect to select the global minimum with probability 1.

We summarize the results from example 4 in the following proposition.

Proposition 6.6. *Let $n = 4$ and the knowledge network be a circle. Then $\hat{m}$ as defined in example 4 is an equilibrium in which the principal does not select the global minimum with probability 1 for some t.*

Example 4 raises the important question for which network architectures the mechanism π achieves full implementation. This question is not answered in this paper. Our conjecture is that π guarantees full implementation for every network architecture in which one agent has full information. This means that there exists one agent who is linked to every other agent.

7 Conclusion

In this paper, we introduce a model in which a principal has to assign a prize to one agent out of a set of heterogeneously valued agents. The principal is uninformed in the sense that she only knows the distribution of values across agents. Every agent would like to get the prize and she exactly knows her own value for the principal. Moreover, an agent might have knowledge about other competing agents. The distribution of information across agents is described by their knowledge network. In addition to her own value an agent exactly knows the values of other agents she is linked to in the network. A link signifies familiarity between two agents, e.g. two researcher who are coauthors have knowledge about each other's abilities.

We propose a mechanism for the principal to identify the best agent in such an environment, when 1) agents send a private costless messages to the principal containing statements about their own value (application) and about values of other agents they know (references), 2) agents can lie about true values only to a commonly known extent, and 3) prefer to tell the truth if they cannot increase their expected probability of being selected by lying. The mechanism specifies a probability of being selected for every agent for any possible message profile. Intuitively, the proposed mechanism encourages agents who have a positive expectation of being selected and who cannot prove better than all their neighbors to exaggerate positively about themselves. Moreover, truthful references are necessary to distinguish between different agents who are all close to the ideal. We show that if every agent has at least one neighbor, this mechanism induces a static Bayesian game among agents for which there exists an equilibrium in which the principal selects the best agent with certainty. In this equilibrium, agents who have a positive expectation of being the best agent prove that they are better than all their neighbors or lie positively about themselves and negatively about their neighbors. Agents who know that they are not the best agent say the truth. Moreover, we show that if the knowledge network is complete or a star, the mechanism guarantees that every equilibrium is such that the best agent is identified with certainty. However, we also provide an example of a knowledge network for which the mechanism does not achieve full implementation. A characterization of all knowledge networks for which full implementation is attained and an analysis of which properties determine whether full implementation on a given network is possible or not would be interesting to pursue in future work.

The fact that in our model references of competitors are valuable to distinguish between agents close to the ideal suggests that also in real world allocation problems it might be helpful to request statements from applicants about their competitors to identify the best applicant. Of course, our model speaks to some real world allocation problems more than to others. For example, modeling an agent's value for the principal as a her distance to the ideal is suitable when an agent's value is measured on a continuous space and the ideal point is known. In the real world this is, for example, the case when an individual's abilities and qualifications are summarized in a score. Such scores are sometimes used by committees to evaluate and rank candidates for scholarships or other rewards. We would expect limits to manipulating the "true" score, which are reflected in the b of our model. Usually points attributed to an individual must be well reasoned, but of course it is possible to either emphasize the positive or negative aspects and by that distort an individual's true value.

Our assumption that b is common knowledge and a candidate perfectly knows her own value and the one of her "neighbors" does not account for existing uncertainties and partial knowledge in the real world. Such uncertainty about an agent's own value and about her neighbors' values is an important extension of the model to investigate. Moreover, the knowledge network is probably common knowledge in only few real world situations. This leads us to another interesting alteration of the model: to allow the existence of a link to be private knowledge and agents to be silent about neighbors. Agents might want to conceal that they have knowledge about another neighbor or might not want to make a statement about one of her neighbor if it improves their own expected probability of being selected. It is unclear if with such extensions we can still find a mechanism to identify the best agent with certainty.

References

Acemoglu, D., Bimpikis, K. and Ozdaglar, A. (2014), 'Dynamics of information exchange in endogenous social networks', *Theoretical Economics* **9**(1), 41–97.

Ben-Porath, E. and Lipman, B. L. (2012), 'Implementation with partial provability', *Journal of Economic Theory* **147**(5), 1689–1724.

Bloch, F., Demange, G. and Kranton, R. (2014), 'Rumors and social networks', *PSE Working Papers n2014-15, halshs-00966234* .

Calvó-Armengol, A., Martí, J. and Prat, A. (2015), 'Communication and influence', *Theoretical Economics* **10**(2), 649–690.

Currarini, S. and Feri, F. (2015), 'Information sharing networks in linear quadratic games', *International Journal of Game Theory* **44**(3), 701–732.

Deneckere, R. and Severinov, S. (2008), 'Mechanism design with partial state verifiability', *Games and Economic Behavior* **64**(2), 487–513.

Förster, M. (2015), 'Strategic communication under persuasion bias in social networks', *Available at SSRN 2610768* .
URL: *https://ssrn.com/abstract=2610768*

Galeotti, A., Ghiglino, C. and Squintani, F. (2013), 'Strategic information transmission networks', *Journal of Economic Theory* **148**(5), 1751–1769.

Garcia, D. (2012), 'Communication and information acquisition in networks', *MPRA Paper No. 55481* .
URL: *http://mpra.ub.uni-muenchen.de/55481/*

Glazer, J. and Rubinstein, A. (2001), 'Debates and decisions: On a rationale of argumentation rules', *Games and Economic Behavior* **36**(2), 158–173.

Hagenbach, J. and Koessler, F. (2010), 'Strategic communication networks', *The Review of Economic Studies* **77**(3), 1072–1099.

Kartik, N. and Tercieux, O. (2012), 'Implementation with evidence', *Theoretical Economics* **7**(2), 323–355.

Lipman, B. L. and Seppi, D. J. (1995), 'Robust inference in communication games with partial provability', *Journal of Economic Theory* **66**(2), 370–405.

Patty, J. W. and Penn, E. M. (2014), 'Sequential decision making and information aggregation in small networks', *Political Science Research and Methods* **2**(02), 243–271.

Renou, L. and Tomala, T. (2012), 'Mechanism design and communication networks', *Theoretical Economics* **7**(3), 489–533.

Wolitzky, A. (2015), 'Communication with tokens in repeated games on networks', *Theoretical Economics* **10**(1), 67–101.