Self-Ratings and Peer Review

Job Market Paper

Leonie Baumann[*,†]

11th November 2018

Abstract A principal has to allocate a prize without monetary transfers. She wants to give it to the most valuable agent but does not know any agent's value. Agents' information is described by a network: Each agent knows her own value and the values of her neighbors. Given a principal's prize allocation rule, agents compete for the prize and send messages about themselves (*application*) and about their neighbors (*references*) to the principal. They can lie, but only to a certain extent. Can full implementation be obtained? This means, is there a prize allocation rule such that the best agent gets the prize in every equilibrium? Bayesian-monotonicity and the revelation principle fail in this setup. I propose a mechanism which allocates the prize as a function of "best" applications and "worst" references. This mechanism fully implements the principal's objective if the network is complete. In environments where agents only lie if it increases their chances of winning, an extended version of the mechanism fully implements the principal's objective for a larger class of networks.

Keywords network, mechanism design, prize allocation, full implementation, hard evidence, partial honesty

JEL Classification Codes C72, D82, D83

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

[†]A previous version of this paper is "Identifying the best agent in a network" (Baumann, 2017). I am especially grateful to Anke Gerber and Ariel Rubinstein for comments and discussion. I also thank Francis Bloch, Marcin Dziubiński, Matt Elliott, Ben Golub, Sanjeev Goyal, Matt Jackson, Claudia Herresthal, Mikhail Safronov, seminar participants in Berlin, Cambridge, Málaga, Northwestern Kellogg, Paris School of Economics, Stanford and Warsaw, and participants in the 3rd Annual Conference on Network Science and Economics, CTN Annual Workshop 2017, ESEM 2017, VfS Annual Conference 2017, BiNoMa Workshop 2018, Barcelona GSE Summer Forum: Networks Workshop 2018. While working on this paper, I received financial support from the University of Hamburg and the Cambridge INET-Institute.

1 Introduction

Consider a setting in which a principal has to allocate a prize without monetary transfers. She wants to give it to the most valuable of all agents but does not know any agent's value. All agents want the prize and their information is described by a network: Each agent knows her own value and the values of agents whom she is linked to. The distribution of values is common knowledge, and there are a minimum and a maximum possible value any agent can have. Given a principal's prize allocation rule, agents compete for the prize and send messages about themselves (*applications*) and about their network neighbors (*references*) to the principal. Agents can lie but only to a certain extent. Given a network, can full implementation be obtained? This means, is there a prize allocation rule such that the best agent gets the prize in every equilibrium? This paper develops a mechanism which always allocates the prize to the best agent in every equilibrium if every agent knows every other agent (complete knowledge network). An extended version of the mechanism always allocates the prize to the best agent in every equilibrium for a larger class of networks, with the additional assumption that agents only lie if lying increases their chances of winning.

The setting resembles several economic environments. An employer has to select an applicant, a manager has to decide whom to promote, or a committee has to assign the prize to the best researcher. Often this decision is based on statements about candidates' qualifications only and candidate-specific monetary transfers to the principal are precluded. Candidates for the prize can have information about each other, for example, two individuals who have worked together know each others' abilities. There are a minimum and a maximum possible value for individuals' qualifications if, for example, they are rated according to scales with an upper and a lower bound. In case the principal is informed about the knowledge network, she can solicit an application and references about neighbors from each candidate. Candidates face a limit to lying when there are psychological or physical costs from lying, or when applications and references have to be supported with evidence. Evidence is often not fully conclusive from the principal's point of view, for example, individuals can distort and frame evidence in different ways. A limit to lying in the context of mechanism design has first been analyzed by Green and Laffont (1986) and is a case of what has been termed "hard evidence" or "partially verifiable information" by the literature.

An important monotonicity property, widely used for full implementation, is not satisfied in our model. Maskin-monotonicity, the respective condition for the complete knowledge network, and Bayesian-monotonicity, the respective condition for any incomplete knowledge network, fail because every agent prefers a higher over a lower probability of getting the prize for any profile of agents' values. As the monotonicity condition does not hold, it follows from Maskin (1999) and Jackson (1991) that there would be no fully implementing mechanism in our setting, if agents did not face a limit to lying.

The limit to lying leads to a failure of the revelation principle in our model because "worse" agents cannot fully imitate "better" agents (see section 4.2). The failure of the revelation principle implies that if there is full implementation, then it is not necessarily achieved via truthful revelation. Indeed, the mechanisms which I use for full implementation induce equilibria in which agents lie.

I first show that a mechanism which allocates the prize as a function of applications only does not achieve full implementation for any knowledge network: If all agents have a sufficiently high value, then all claim to have the maximum possible value and the principal cannot distinguish between them. Thus, full implementation in our setting can only be obtained if every agent has at least one neighbor because there must be at least one reference for every agent.

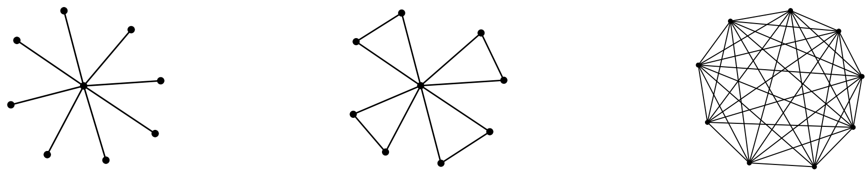
Similarly, a mechanism which allocates the prize as a function of references only does not achieve full implementation for any knowledge network: If all agents have a sufficiently low value, then there is an equilibrium in which every agent depicts all neighbors as having the minimum possible value and the principal cannot distinguish between agents.

I then propose the simple mechanism π^{so} which allocates the prize as a function of "best" applications and "worst" references. The mechanism awards the prize to the agent who sends the best application and the worst reference about her is weakly better than the worst reference about any agent who sends a best application. Full implementation is achieved if every agent knows every other agent (complete knowledge network): The best agent can send the best possible application. Furthermore, if many agents can claim to have the maximum possible value, then the best agent can send worse references about every other agent than any agent can send about the best agent. This is because of the limit to lying.

For every incomplete knowledge network, the mechanism π^{so} guarantees the existence of a "salient" dominant strategy equilibrium such that the best agent gets the prize. In this equilibrium, every agent fully exaggerates – positively about herself and negatively about her neighbors. However, there also exist equilibria such that the best agent does not get the prize for some profiles of agents' values: Suppose exactly two agents claim to have the maximum possible value and they are not linked to each other. Then the references from their neighbors determine which of the two gets the prize. For some value profiles, these neighbors are certain not to get the prize because they cannot send a weakly better application. In this case, the neighbors are indifferent between all references they can send and the best agent does not get the prize for some equilibria.

Consequently, I study the model with the additional assumption of *partial honesty*. Partial honesty means that agents only lie if lying increases their expected probability of winning. This concept has first been introduced to the implementation literature by Dutta and Sen (2012). With partial honesty, mechanism π^{soh}, an extended version of π^{so}, fully implements the principal's objective for a larger class of networks. If the network is connected and for every agent and each of her neighbors it is true that 1) the agent is linked to all the neighbors of her neighbor, or 2) the neighbor is linked to every other agent, then π^{soh} fully implements with partially honest agents. Condition 1) means that the agent knows all the values which her neighbor knows, and 2) means that the neighbor knows every value. Networks which satisfy these conditions are for example the star network, the complete network, and the Dutch windmill, as depicted in the following figure.

I find that the property of full implementation with π^{soh} and partially



(a) Star (b) Dutch Windmill (c) Complete Network

honest agents is non-monotonic in the number of links in the network. There is a network which is a supergraph of the star and a subgraph of the complete network such that π^{soh} is not fully implementing: for one equilibrium, the best agent and another agent who is not linked to the best agent both get the prize with probability .5. Both of them claim to have the maximum possible value and they receive the same untruthful reference. Each agent who sends an untruthful reference about the awarded agents expects herself to be the falsely awarded agent such that lying is optimal for her. Hence, more links, that is more information and more messages, are not always beneficial for full implementation.

Finally, I discuss the effect of noise in the communication between the agents and the principal on full implementation. If the principal receives agents' messages with a random noise term, then there is a unique equilibrium with mechanism π^{so} in which every agent fully exaggerates – positively about herself and negatively about her neighbors. This leads to full implementation in expectation for every network with mechanism π^{so}.

I describe two applications of the model: employee performance evaluation (*360 degree feedback*) and peer review processes in academia. Our results relate to the interpretation of self- and other-ratings by an employer and to the evaluation of proposals/papers and reports by an academic principal (editor, conference organizer, or funding institution), given that the principal wants to identify the best employee/researcher and can rely on a limit to lying.

This paper contributes to the following strands of literature. First, consider the literature on mechanism design for allocation and persuasion problems, and on implementation of social choice functions with one principal, multiple agents, and partially verifiable information. Determining agents' information and message space by a network is new to this literature. Lipman and Seppi (1995), Glazer and Rubinstein (2001), Ben-Porath and Lipman (2012), and Kartik and Tercieux (2012) assume that all agents are fully informed or have the same information about the state of the world. A network allows for differentially informed agents. In Deneckere and Severinov (2008) agents have partially verifiable information about their own type but not about other agents' types. The network structure of information in our setting means that agents have partially verifiable information about themselves but also about their neighbors. The latter is necessary for full implementation in our setup. Koessler and Perez-Richet (2017) show for a general setup that a social choice is partially implementable if every type has a message available which no other type wants to or is able to imitate. They call this the evidence base condition. This condition is not satisfied in our setting.

Only a few papers in economics combine mechanism design and network theory. Renou and Tomala (2012) study the implementation of social choice functions for different communication networks between agents and a mechanism designer. Dziubiński et al. (2016) analyze optimal network protection against an attacker when the network is unknown to the defender and imperfectly to the nodes in the network. The paper closest to ours is recent subsequent work by Bloch and Olckers (2018). In their setting, the principal's objective is to extract a complete ordinal ranking of agents in a social network. They focus on truthful mechanisms and analyze the network structures for which there is an equilibrium such that the outcome corresponds to the principal's objective.

Recently, a literature on peer review systems in academia where reviewers also compete for the prize has developed in computer science. Some of the proposed peer review systems have been implemented for conferences (Nierstrasz, 2000) or grant allocations by the National Science Foundation (Merrifield and Saari, 2009). These systems rely on references only and do not award the prize to the best agent in all equilibria. Sometimes, even the existence of one desirable equilibrium is not guaranteed. Consequently, Kurokawa et al. (2015) and Aziz et al. (2016) have focused on the design of a references-based mechanism without a limit to lying which guarantees one desirable, truthful equilibrium. Such a mechanism cannot fully implement in our setting. Moreover, this literature assumes that an agent's value is determined by other agents' potentially heterogeneous preferences over the set of agents, and not, as in this paper, by a principal's preference ordering over agents or by an unambiguous true state of the world.

This paper proceeds as follows. In Section 2, I introduce the model. Section 3 discusses employee performance evaluation and peer review processes in academia as two applications. Section 4 shows why Maskin-/Bayesian-monotonicity and the revelation principle fail in our setup, and that a mechanism which uses either applications or references only does not fully implement. Section 5 presents the main analysis. In Subsection 5.1, I introduce mechanism π^{so} which relies on both applications and references. I show that π^{so} fully implements if the network is complete. In Subsection 5.2, mechanism π^{soh}, an extension of π^{so}, is proposed. I characterize a class of networks for which π^{soh} fully implements if agents are partially honest. Subsection 5.3 discusses that π^{so} fully implements in expectation in all networks if communication is noisy. Finally, Section 6 concludes.

2 The Model

A principal has to assign an indivisible prize to one agent out of a set of agents $N = \{1, ..., n\}$ where $n \geq 3$. Agent i gets utility $v_i > 0$ from receiving the prize and 0 from not receiving it.

Agents differ in their suitability to receive the prize. There exists a combination of characteristics which is the ideal match for the prize. This ideal is common knowledge. Agent i's suitability to receive the prize is measured by her distance d_i to the ideal. We assume that there is a maximum possible distance to the ideal and w.l.o.g we normalize it to 1. Then $d_i \in [0, 1]$ for all $i \in N$ with 1 being the worst possible fit and 0 the ideal match. We assume that each d_i is independently and identically drawn from a continuous full support distribution without atoms over $[0, 1]$.[1] The distribution is common

[1]This way of modeling heterogeneity captures n-dimensional characteristics. Consider any compact subset $C \subset \mathbb{R}^n$. Let $c^* \in C$ be the ideal match for the principal. Agent

knowledge. The principal's utility is strictly decreasing in the distance of the agent who receives the prize. Hence, for any realization of distances, the best possible outcome for the principal is to assign the prize to agent i with $d_i = \min_{k \in N} d_k$ whom we refer to as the *global minimum* g. With probability 1, a realization of distances is such that $d_i \neq d_j$ for all i and $j \neq i$ and we restrict the analysis to such cases.

The principal does not know the distance of any agent but only the distribution. Every agent exactly knows her own distance and the distances of certain other agents as specified by a graph which is common knowledge. The graph is undirected and given by a set of links L among the agents. Links signify familiarity. If link $ij \in L$, then agent i and agent j exactly know each other's distance. The set of neighbors of agent i is $N_i := \{j \,|ij \in L\}$. Regarding the distances of agents who are not her neighbors, agent i only knows the distribution. Agent i's type is $\theta_i = \left(d_i, (d_j)_{j \in N_i}\right)$ and summarizes her knowledge. Agent i's type space is $\Theta_i = [0,1]^{|N_i|+1}$. A type profile is denoted by θ and the set of all feasible type profiles by Θ. If a profile or set contains all elements except for agent i's, we conventionally mark the profile or set with subscript $-i$. Given type realization θ, agent i's posterior that $\theta_{-i} \in \Theta'_{-i} \subset \Theta_{-i}$ is the conditional probability $p(\Theta'_{-i}|\theta_i)$ derived from the distribution of distances. We assume in the following that every agent has at least one neighbor. Without this assumption, the principal could never identify the global minimum with probability 1. The reason will become obvious in Subsection 4.2. Note that this assumption does not require the graph to be connected.

The principal designs a mechanism with the objective to "always" identify the global minimum. Any mechanism is a pair (M, π) where $M = \prod_{i \in N} M_i$ is a set of message profiles and π specifies an outcome for every message profile $m \in M$. The message space for agent i is $M_i = \Theta_i$. Thus any message $m_i \in M_i$ of agent i is such that $m_i = \left(m_{ii}, (m_{ij})_{j \in N_i}\right)$ with $m_{ii}, m_{ij} \in [0,1]$ for all $j \in N_i$. Agent i makes statement m_{ii} about her own distance which we call

i's characteristic is $c_i \in C$, iid from a continuous, full support distribution without atoms over C. Then agent i has a Euclidean distance d_i to c^* and we normalize the maximum possible distance to 1.

her *application* and statement m_{ij} about the distance of her neighbor $j \in N_i$ which we call her *reference* about j. Since M is fixed for all mechanisms the principal considers, we simply refer to π as the mechanism. We assume that the principal can choose any $\pi : M \to [0,1]^n$ with $\sum_{i \in N} \pi_i(m) = 1$ for all $m \in M$. Outcome $\pi(m)$ is a probability distribution over N with $\pi_i(m)$ being the probability that agent i receives the prize if the message profile is m. Probabilities sum up to 1 for every m because the principal has to assign the prize.[2] The principal does not use transfers.

Any mechanism π induces the following static Bayesian game $\Gamma(\pi)$ among the agents. Agent i's action set at θ_i is message set $M_i(\theta_i) \subset M_i$. Message $m_i \in M_i(\theta_i)$ if and only if $m_{ik} \in [\max\{0, d_k - b\}, \min\{d_k + b, 1\}]$ for all $m_{ik} \in m_i$ with exogenous and commonly known $b \in (0, \frac{1}{2})$. This means agents can lie about each true distance maximally by $\pm b$ and an agent's action set varies with her type. Such type-dependent action sets are a setting of what has been termed "hard evidence" in the literature.

The limit to lying is an abstraction of psychological or physical lying costs or of inconclusive evidence: "Cheating" within certain bounds but not beyond might be morally acceptable. Framing inconclusive evidence more positively or negatively than consistent with the truth might be possible but only up to a certain limit. Different types might have access to some same pieces of evidence, as is the case with degree certificates. Both an applied economist and a theorist have a PhD in Economics.

Agent i's strategy is a function $\hat{m}_i$ that specifies $m_i \in M_i(\theta_i)$ for every $\theta_i \in \Theta_i$. We restrict ourselves to pure strategies. A strategy profile is denoted by $\hat{m}$. Let $\hat{m}_{-i}(\theta_{-i})$ be the strategies of all agents other than i at θ_{-i}.

Given $\hat{m}_{-i}$, agent i's expected utility at θ_i from choosing $m_i \in M(\theta_i)$ is

$$U_i(m_i, \hat{m}_{-i}|\theta_i) = v_i \int_{\theta_{-i}} \pi_i(m_i, \hat{m}_{-i}(\theta_{-i}))\ dp(\theta_{-i}|\theta_i).$$

A strategy profile $\hat{m}$ is a Bayesian Nash equilibrium of $\Gamma(\pi)$, if for all i,

[2] This assumption implies that this is not a separable environment as defined by Jackson et al. (1994).

$m_i \in M_i(\theta_i)$, and θ_i

$$U_i(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i) \geq U_i(m_i, \hat{m}_{-i}|\theta_i).$$

It will be useful to denote agent i's expected probability of receiving the prize at θ_i when choosing $m_i \in M(\theta_i)$ given $\hat{m}_{-i}$ as

$$\Pi_i(m_i, \hat{m}_{-i}|\theta_i) = \int_{\theta_{-i}} \pi_i(m_i, \hat{m}_{-i}(\theta_{-i}))\ dp(\theta_{-i}|\theta_i).$$

Clearly, a strategy profile $\hat{m}$ is a Bayesian Nash equilibrium of $\Gamma(\pi)$ if and only if for all i, $m_i \in M_i(\theta_i)$, and θ_i

$$\Pi_i(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i) \geq \Pi_i(m_i, \hat{m}_{-i}|\theta_i).$$

The goal of the principal is to design π such that the global minimum gets the prize with probability 1 given graph L. We say that mechanism π *fully implements* the principal's objective in L if every equilibrium $\hat{m}$ of $\Gamma(\pi)$ is such that $\pi_g(\hat{m}(\theta)) = 1$ with probability 1. We say that mechanism π *partially implements* the principal's objective in L if there exists an equilibrium $\hat{m}$ of $\Gamma(\pi)$ such that $\pi_g(\hat{m}(\theta)) = 1$ with probability 1. Note that if mechanism π is partially implementing, then there can exist other equilibria of $\Gamma(\pi)$ in which the global minimum does not get the prize with probability 1.

3 Two Applications

We map the model to two applications. The first application is *360 degree feedback*, a method widely used by companies to evaluate employee performance. The second one is the peer review process in academia.

Example 1. 360 degree feedback.

360 degree feedback describes a process in which information about an employee is gathered from different sources including the employee himself and, for example, co-workers, subordinates, and managers. Nowadays, the information from each source is usually collected through electronic questionnaires which include mostly numerical ratings and some open questions and written comments.

The method is a popular tool for employee evaluation and development. A 2013 survey (3D Group, 2016) conducted among 112 companies, including Boeing, Monsanto, PepsiCo, FedEx and Dell, reports that 34% of the surveyed companies use 360 degree feedback for pay and promotion decisions. 79% of the surveyed companies expect their use of 360 degree feedback to remain constant or to increase. Roughly 65% of the companies hire an external service provider to implement the 360 degree feedback. Established consultancies and software providers specialized in human capital management like Cornerstone OnDemand ($482 million revenue and 1891 employees in 2017 (Cornerstone OnDemand, 2018)) and SABA (1000 employees (Saba Software, 2018)), as well as small specialized firms like 3D Group (3D Group, 2018) or ETS (Expert Training Systems, 2018) offer 360 degree feedback products. The natural concern for vendors and companies using 360 degree feedback is how to correctly interpret the questionnaires and how "reliable" the ratings are.

The following precise situation is an example for our model. A human resource manager has to pick one employee as the team leader, for a promotion or as employee of the month, all of which is desirable for each employee. To identify the "best" employee, the manager asks every candidate i for a self-rating and for a rating of the other candidates who i has worked with. Ratings are confidential and only seen by the manager. A limit to lying arises if the manager asks for justification of ratings.

Example 2. Peer-review in academia.

Peer review is essential to many selection processes in academia. We explain how the review processes for conferences and journals relate to our model.

Consider the peer review processes in the run-up for top conferences in the field of artificial intelligence (e.g. AAAI, IJCAI, AAMAS) or at the intersection of economics and computation (e.g. ACM EC).[3] The principal in this context is the conference organizer who awards different "prizes" like slots for presentation and the best paper award. A double blind peer review process determines which of the submitted papers get accepted and which paper gets the best paper award. Within the framework of our model, a paper submission is an application. Each submitted paper is reviewed by 4-5 referees who usually have submitted a paper themselves and are thus competing with other submissions. Each referee has to grade a paper in different categories like readability, novelty, importance of the contribution, technical quality, and also conclude a general grade which is often numerical. All grades in a report have to be justified. Finally, all papers are ranked according to their scores and whether a paper gets selected depends on its score rank.

The review process for journal publications does not map as closely to the model as the one for conferences, but shows the same mechanisms and incentives. An author submits a paper (an application) to a journal for publication and an editor (the principal) selects peer reviewers who write reports about the paper. Peer reviewers who work on topics related to the paper are competing with the author for publication. Reports have to be well grounded and justified. Some journals require each referee to assign an overall numerical grade to the paper.

Within our model, the "true distance" of a paper corresponds to the true quality of the research it presents. The paper is the "application". The "application" and the "references" are not necessarily truthful. An author can try to hide critical assumptions or oversell the contribution. Referees can criticize assumptions and results in an exaggerated way. However, there is a natural limit to lying. All claims have to be supported with some evidence. Results in the paper need to be supported by proofs, and criticism by a referee has to be supported by examples.

[3] I thank Ben Golub and Marcin Dziubiński for insightful discussions on this topic.

4 Unsuccessful Mechanisms

This section first shows that some popular and intuitive mechanisms fail to fully implement the principal's objective. First, mechanisms which rely on Maskin-/Bayesian-monotonicity fail, because Maskin-/Bayesian-monotonicity is violated in our setup (Subsection 4.1). This implies that full implementation would be impossible in our setup without hard evidence (Jackson, 1991; Maskin, 1999). Thus, any successful mechanism must exploit the hard evidence. The presence of hard evidence leads to a failure of the revelation principle such that truthful implementation might not be possible.

Consequently, we study mechanisms which make use of the hard evidence and which implement via non-truthful equilibria. Mechanisms which rely on either applications or references only still fail to implement (Subsections 4.2 and 4.3). However, they fail for different type realizations such we continue to show in Section 5 that mechanisms which combine applications and references are successful in full implementation.

4.1 Mechanisms relying on Maskin- or Bayesian-Monotonicity

Maskin- or Bayesian-monotonicity guarantee in many setups that a fully implementing mechanism exists. Popular mechanisms which rely on Maskin- or Bayesian monotonicity for full implementation are "consensus" mechanisms. In such mechanisms, if there is a consensus, this is $m_{ji} = m_{ki}$ for all i, j, k in our setup, then the principal chooses the outcome according to the consensus. A concern with such mechanisms there might be equilibria where the consensus is not truthful such that the principal picks the "wrong" outcome. Maskin-/Bayesian-monotonicity is used to rule out undesired equilibria in settings of complete/incomplete information among agents. For Maskin- and Bayesian-monotonicity to be satisfied, agents' preferences over outcomes must change across different type realizations. In our setup, no agent experiences a preference reversal over outcomes for any two type realizations. Every agent always strictly prefers a higher over a lower expected probability of winning. Thus, Maskin- and Bayesian-monotonicity fail and cannot be used to rule out undesirable consensus equilibria. The failure of Maskin- and Bayesian-monotonicity implies that if our setup did not feature hard evidence (the limit to lying), then full implementation would be impossible (Jackson, 1991; Maskin, 1999). Kartik and Tercieux (2012) show that hard evidence can recover full implementation for complete information settings among agents, if Maskin-monotonicity fails. For incomplete information settings among agents, the conditions under which hard evidence recovers full implementation, if Bayesian monotonicity fails, are less clear. Koessler and Perez-Richet (2017) take some steps in this direction.

The hard evidence in our setup leads to a failure of the revelation principle, as will be shown in the next subsection. This issue has first been pointed out by Green and Laffont (1986). Thus, if full implementation is possible, then it is not necessarily achievable via truthful revelation mechanism. In the following, we focus on mechanisms which exploit the hard evidence property and which do not have truthful equilibria.

4.2 Applications only

Suppose the principal uses a mechanism π which exploits the hard evidence but only relies on applications and disregards all references. Then the principal does not identify the global minimum with probability 1 for type realizations where many agents have distances less than b. We define the following applications-only mechanism.

Definition 1. ***Applications-Only Mechanism*** π^s

For any $m \in M$,
choose $B_1(m) \subseteq N$ *such that* $i \in B_1(m)$ *if and only if* $m_{ii} = \min_{k \in N} m_{kk}$.
Let $\pi_i^s(m) = \frac{1}{|B_1(m)|}$ *for all* $i \in B_1(m)$ *and* $\pi_i^s(m) = 0$ *for all* $i \notin B_1(m)$.

In words, for any m, the principal identifies all agents who send the best application and assigns the prize with equal probability to one of them. Observe that $B_1(m)$ is never empty.

Lemma 4.1. *Let $D(\theta) = \{i \mid d_i \leq b\}$ for all θ. For any L, every equilibrium $\hat{m}$ of $\Gamma(\pi^s)$ is such that with probability 1*

$\pi^s_g(\hat{m}(\theta)) = 1$ if $d_g > b$, and

$\pi^s_i(\hat{m}(\theta)) = \frac{1}{|D(\theta)|}$ for all $i \in D(\theta)$ if $d_g \leq b$.

Proof. Assume $\hat{m}$ is an equilibrium. Suppose that g has $d_g > b$ and expects with positive probability that $\pi^s_g(\hat{m}(\theta)) < 1$. Then g can deviate to m'_g with $m'_{gg} = d_g - b$ such that $\pi^s_g(m'_g, \hat{m}_{-g}(\theta_{-g})) = 1$. For any other case which g expects with positive probability and for which $g \in B_1(\hat{m}(\theta))$, $B_1(m'_g, \hat{m}_{-g}(\theta_{-g})) \subset B_1(\hat{m}(\theta))$ such that g would not do worse. Thus,

$\Pi(m'_g, \hat{m}_{-g}|\theta_g) > \Pi(\hat{m}_g(\theta_g), \hat{m}_{-g}|\theta_g)$ and $\hat{m}$ is not an equilibrium.

Suppose g has $d_g \leq b$ and agent i with $d_i \leq b$ expects with positive probability that $\pi^s_i(\hat{m}(\theta)) < \frac{1}{|D(\theta)|}$. Then i can deviate to m'_i with $m'_{ii} = 0$ such that $\pi^s_i(m'_i, \hat{m}_{-i}(\theta_{-i})) \geq \frac{1}{|D(\theta)|}$. For any other case which i expects with positive probability and for which $i \in B_1(\hat{m}(\theta))$, $B_1(m'_i, \hat{m}_{-i}(\theta_{-i})) \subset B_1(\hat{m}(\theta))$. Thus $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ and $\hat{m}$ is not an equilibrium.

Suppose g has $d_g \leq b$ and agent i with $d_i \leq b$ expects with positive probability that $\pi^s_i(\hat{m}(\theta)) > \frac{1}{|D(\theta)|}$. Then some agent j with $d_j \leq b$ would expect with positive probability that $\pi^s_j(\hat{m}(\theta)) < \frac{1}{|D(\theta)|}$ which is a contradiction. □

At this point, it is convenient to explain why the revelation principle does not hold in our model. The revelation principle would claim for our setup that if $\hat{m}$ is an equilibrium of $\Gamma(\pi^s)$, then truthful revelation is an equilibrium of $\Gamma(\pi^s \circ \hat{m})$ where the principal first executes the equilibrium strategies $\hat{m}$ for the agents after they have announced their types and, second, the principal executes π^s. The revelation principle fails because better types have actions available which worse types do not have available and which worse types prefer over any of their available actions. This illustrated by the following example.

Suppose $n = 3$, L complete and $b = .2$. First observe that any strategy profile $\hat{m}$ such that $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ for all θ_i and all i is an equilibrium of $\Gamma(\pi^s)$. Assume now the principal uses mechanism $\pi^s \circ \hat{m}$, the distance realization is $d_1 = .7, d_2 = .8, d_3 = .9$ and all agents report their types truthfully. Then agent 2 has an incentive to deviate to m'_2 with

$m'_{22} = .6$ because $\hat{m}_{22}(m'_2) = .6 - .2 < \hat{m}_{11}(\theta_1) = .7 - .2 < \hat{m}_{22}(\theta_2) = .8 - .2$ such that $\pi^s_2(\hat{m}(m'_2, \theta_{-2})) = 1 > \pi^s_2(\hat{m}(\theta)) = 0$. Thus, truthful revelation is not an equilibrium.[4]

4.3 References only

Suppose next the principal disregards all applications and only consults references. Then the principal only identifies the global minimum with probability $\frac{1}{n}$ for any dominant strategy equilibrium and any graph, in case all agents are sufficiently bad ($d_g \geq 1 - b$). If the graph is complete, then every equilibrium has this property. We define the following references-only mechanism. Let $\bar{r}_i(m) = \max_{j \in N_i} m_{ji}$, this means $\bar{r}_i$ is the worst (maximum) reference about i given m.

Definition 2. ***References-Only Mechanism*** π^o

For any $m \in M$,
choose $B_2(m) \subseteq N$ *such that* $i \in B_2(m)$ *if and only if* $\bar{r}_i = \min_{k \in N} \bar{r}_k$.
Let $\pi^o_i(m) = \frac{1}{|B_2(m)|}$ *for all* $i \in B_2(m)$ *and* $\pi^o_i(m) = 0$ *for all* $i \notin B_2(m)$.

Thus, for any m, the principal identifies all agents who receive the least bad worst reference, in other words, the min-max reference, and assigns the prize with equal probability to one of them. Observe that $B_2(m)$ is never empty.

For our next result, we define a dominant message and dominant strategy.

Let $M_{-i}(\Theta'_{-i}) := \bigcup_{\theta_{-i} \in \Theta'_{-i}} M_{-i}(\theta_{-i})$ where $\Theta'_{-i} \subset \Theta_{-i}$ and $M_{-i}(\theta_{-i})$ is set the of message profiles which agents $j \neq i$ can send if they have types θ_{-i}.

[4] Strausz (2016) argues that the validity of the revelation principle can be recovered in settings with hard evidence if the mechanism can be adjusted such that an agent can claim to be any type, and the outcome function then specifies which evidence has to be submitted and how the prize is assigned. The separation of type and evidence announcement by making the evidence requirement part of the outcome function is is a reformulation of the problem which shifts the untruthfulness from the type announcements to the evidence requirement in the outcome function. That is if the revelation principle fails in the original setup, then the sufficiency condition for the existence of an implementing mechanism with truthful evidence requirement fails. Thus, the difficulty from finding an implementing mechanism with untruthful equilibria is translated into the difficulty of designing the “untruthful” outcome function. The problem is equivalent and forfeits the simplifying purpose of the revelation principle.

Definition 3. ***Dominance***

A message $m_i \in M(\theta_i)$ is dominant at θ_i if $\Pi_i(m_i, m_{-i}|\theta_i) \geq \Pi_i(m_i', m_{-i}|\theta_i)$ for all $m_i' \in M(\theta_i)$, all $m_{-i} \in M_{-i}(\Theta'_{-i})$ and all $\Theta'_{-i} \subset \Theta_{-i}$ for which $p(\Theta'_{-i}|\theta_i) > 0$.

A strategy $\hat{m}_i$ is dominant, if $\hat{m}_i(\theta_i)$ is a dominant message at θ_i for all θ_i.

A dominant strategy equilibrium is a Bayesian Nash equilibrium in which every agent agent chooses a dominant strategy.

In words, a message $m_i \in M(\theta_i)$ is dominant at θ_i if it maximizes agent i's expected utility for all messages agents $j \neq i$ can send for any type profile which agent i expects with positive probability to be the true type realization. A strategy is dominant if it only consists of dominant messages.

Lemma 4.2. *For any L, every dominant strategy equilibrium $\hat{m}$ of $\Gamma(\pi^o)$ is such that $\pi_i^o(\hat{m}(\theta)) = \frac{1}{n}$ for all $i \in N$ if $d_g \geq 1 - b$.*

If L is complete, then every equilibrium $\hat{m}$ of $\Gamma(\pi^o)$ is such that $\pi_g^o(\hat{m}(\theta)) = 1$ if $d_g < 1 - b$, and $\pi_i^o(\hat{m}(\theta)) = \frac{1}{n}$ for all $i \in N$ if $d_g \geq 1 - b$.

Proof. Observe that $d_g \geq 1 - b$ implies $d_i \geq 1 - b$ for all $i \in N$. Consider θ_i such that $d_i, d_j \geq 1 - b$ for all $j \in N_i$. Then no message m_i with $m_{ij} < 1$ for some $j \in N_i$ is dominant at θ_i because if m_{-i} is such that $\bar{r}_j(m_i, m_{-i}) < \bar{r}_k(m_i, m_{-i}) = 1$ for $j \in N_i$ and all $k \neq j$, then $\Pi(m_i', m_{-i}|\theta_i) > \Pi(m_i, m_{-i}|\theta_i)$ for m_i' with $m_{ij}' = 1$ for all $j \in N_i$. Thus every dominant strategy equilibrium $\hat{m}$ is such that $\hat{m}_{ij}(\theta_i) = 1$ for all $j \in N_i$ and all i such that $\pi_i^o(\hat{m}(\theta)) = \frac{1}{n}$ for all i if $d_g \geq 1 - b$.

Next, assume L is complete. Suppose there is an equilibrium $\hat{m}$ such that $\pi_g^o(\hat{m}(\theta)) < 1$ if $d_g < 1 - b$. Then however g can deviate to m_g' with $m_{gj}' = \min\{d_j + b, 1\}$ for all $j \neq g$ such that $\bar{r}_g(m_g', \hat{m}_{-g}(\theta_{-g})) < \bar{r}_k(m_g', \hat{m}_{-g}(\theta_{-g}))$ for all $k \neq g$ and $\pi_g^o(m_g', \hat{m}_{-g}(\theta_{-g})) = 1$.

Suppose there is an equilibrium $\hat{m}$ such that $\pi_i^o(\hat{m}(\theta)) < \frac{1}{n}$ for some $i \in N$ if $d_g \geq 1 - b$. Then however i can deviate to m_i' with $m_{ij}' = 1$ for all $j \neq i$ such that $\bar{r}_i(m_i', \hat{m}_{-i}(\theta_{-i})) \leq \bar{r}_k(m_i', \hat{m}_{-i}(\theta_{-i})) = 1$ for all $k \neq i$ and $\pi_i^o(m_i', \hat{m}_{-i}(\theta_{-i})) \geq \frac{1}{n}$. $\square$

5 Full Implementation with Applications and References

When the principal consults both applications and references and accounts for the limit of lying, then the principal's objective can be fully implemented in certain networks and environments.

In Subsection 5.1, we introduce mechanism π^{so} which is a combination of π^s and π^o and thus takes into account both applications and references. Mechanism π^{so} fully implements in the complete graph. Moreover, for every graph, there is a "salient" dominant strategy equilibrium $\hat{m}$ of $\Gamma(\pi^{so})$ such that $\pi_g^{so}(\hat{m}(\theta)) = 1$ with probability 1. However, for certain graphs there also exist dominant strategy equilibria of $\Gamma(\pi^{so})$ such that the global minimum is identified with probability 0 for some type realizations.

In Subsection 5.2, we show that if agents are partially honest, then mechanism π^{soh} which is an extension of π^{so} fully implements the principal's objective for a larger class of graphs, among which are the star, windmill graphs and the complete graph. We provide sufficient conditions on the graph for full implementation with π^{soh}.

Finally, in Subsection 5.3, we present that if communication is noisy, then π^{so} fully implements the principal's objective in every graph in expectation.

5.1 The Complete Graph

We first define mechanism π^{so} which is a combination of π^s and π^o and thus relies both on applications and references and exploits the limit of lying. This mechanism fully implements in the complete graph. In other graphs, however, only partial implementation is guaranteed.

Definition 4. ***Applications-And-References Mechanism*** π^{so}

For any $m \in M$,

first choose $B_1(m) \subseteq N$ *such that* $i \in B_1(m)$ *if and only if* $m_{ii} = \min_{k \in N} m_{kk}$.

Second choose $B_2(m) \subseteq B_1(m)$ *such that*

if $\min_{k\in N} m_{kk} > 0$, *then* $B_2(m) = B_1(m)$, *and*
if $\min_{k\in N} m_{kk} = 0$, *then* $i \in B_2(m)$ *if and only if* $\bar{r}_i(m) = \min_{k\in B_1(m)} \bar{r}_k(m)$.

Let $\pi_i^{so}(m) = \frac{1}{|B_2(m)|}$ *for all* $i \in B_2(m)$ *and* $\pi_i^{so}(m) = 0$ *for all* $i \notin B_2(m)$.

Thus, for any m, the principal first identifies all agents who send the best application. If the best application is larger than zero, then she assigns the prize with equal probability to one of the agents with the best application. If the best application is zero, then she consults the references and identifies the agent who receives the min-max reference among all agents with the best application. Finally, she assigns the prize with equal probability to agents who send the best application and who receive the min-max reference among the best applying agents. Observe that $B_2(m)$ is never empty.

Proposition 5.1. *Let* L *be complete. Then every equilibrium* $\hat{m}$ *of* $\Gamma(\pi^{so})$ *is such that* $\pi_g^{so}(\hat{m}(\theta)) = 1$ *for all* θ.

The proof of Proposition 5.1 trivially follows from the proofs of Lemmata 4.1 and 4.2. Applications ensure that the global minimum is selected with probability 1, if $d_g > b$, and references ensure that the global minimum is selected with probability 1, if $d_g \leq b$.

The full implementation result for the complete graph is in line with Kartik and Tercieux (2012) who show that there is a mechanism which fully implements the principal's objective, if the hard evidence satisfies a certain monotonicity condition and agents have full information. The evidence structure of our setup satisfies their monotonicity condition if the graph is complete. Kartik and Tercieux (2012) use a mechanism which induces an "integer game" among the agents to prove their result. Such mechanisms involve a more complicated and unnatural message space compared to our setup where agents only choose messages from their type set.

Mechanism π^{so} partially implements for all graphs. For every graph, there is a "salient" dominant strategy equilibrium $\hat{m}$ such that $\pi_g^{so}(\hat{m}(\theta)) = 1$ for all θ. We explain what is meant by "salient" after we have stated and proved the result.

Proposition 5.2. *For any L, strategy profile $\hat{m}$ such that $\hat{m}_{ii}(\theta_i) = \max\{d_i - b, 0\}$ and $\hat{m}_{ij}(\theta_i) = \min\{d_j + b, 1\}$ for all $j \in N_i$, all θ_i and all i is a dominant strategy equilibrium of $\Gamma(\pi^{so})$ for which $\pi_g^{so}(\hat{m}(\theta)) = 1$ for all θ.*

Proof. We first show that m_i with $m_{ii} = \max\{d_i - b, 0\}$ and $m_{ij} = \min\{d_j + b, 1\}$ for all $j \in N_i$ is a dominant message at θ_i for all θ_i and all i. For any $m_i' \in M_i(\theta_i)$ and any m_{-i}, if $i \in B_2(m_i', m_{-i})$, then also $i \in B_2(m_i, m_{-i})$ and $B_2(m_i, m_{-i}) \subseteq B_2(m_i', m_{-i})$. Thus $\pi_i^{so}(m_i, m_{-i}) \geq \pi_i^{so}(m_i', m_{-i})$ and m_i is at least as good as any $m_i' \in M_i(\theta_i)$ for any m_{-i}. Hence, $\hat{m}$ is a dominant strategy equilibrium.

Given $\hat{m}$, $\hat{m}_{gg}(\theta_g) < \hat{m}_{kk}(\theta_k)$ for all $k \neq g$, if $d_g > b$, and $\hat{m}_{gg}(\theta_g) = 0$ and $\bar{r}_g(\hat{m}(\theta)) < \bar{r}_k(\hat{m}(\theta))$ for all $k \neq g$, if $d_g \leq b$. Thus $\pi_g^{so}(\hat{m}(\theta)) = 1$ for all θ. □

The dominant strategy equilibrium $\hat{m}$ defined in Proposition 5.2 is "salient" in the following sense. Each agent's strategy is a simple behavioral rule: always exaggerate to the maximum, positively about oneself and negatively about one's neighbors.

Mechanism π^{so}, however, does not fully implement for all L. For some graphs, there exist dominant strategy equilibria in which the global minimum is identified with probability 0 for some type realizations. The following example illustrates this.

Example 3. Assume $N = \{1, 2, 3\}$ and $L = \{12, 23\}$, thus the graph is a line. Consider $\hat{m}$ such that

- $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ for all θ_i and all i,
- $\hat{m}_{i2}(\theta_i) = \min\{d_2 + b, 1\}$ for all θ_i and $i = 1, 3$,
- $\hat{m}_{21}(\theta_2) = d_3$ and $\hat{m}_{23}(\theta_2) = d_1$, if $d_2 - d_j > 2b$ and $d_j \leq b$ for $j = 1, 3$, and $\hat{m}_{2j}(\theta_2) = \min\{d_j + b, 1\}$ for $j = 1, 3$ otherwise.

We know from the proof of Proposition 5.2 that $\hat{m}_1$ and $\hat{m}_3$ are dominant strategies and that $\hat{m}_2(\theta_2)$ is a dominant message if not both $d_2 - d_j > 2b$

and $d_j \leq b$ for all $j = 1, 3$. We show next that $\hat{m}_2(\theta_2)$ is also a dominant message if both $d_2 - d_j > 2b$ and $d_j \leq b$ for $j = 1, 3$.

If both $d_2 - d_j > 2b$ and $d_j \leq b$ for $j = 1, 3$, then $m_{22} > m_{kk}$ for $k \neq 2$ and thus $\pi_2^{so}(m_2, m_{-2}) = 0$ for any $m_2 \in M_2(\theta_2)$ and any $m_{-2} \in M_{-2}(\theta_{-2})$. Hence, every $m_2 \in M_2(\theta_2)$ is dominant if $d_2 - d_j > 2b$ and $d_j \leq b$ for $j = 1, 3$. Then $\hat{m}$ is a dominant strategy equilibrium.

For all θ where $d_2 - d_j > 2b$ and $d_j \leq b$ for $j = 1, 3$, $\hat{m}_{11}(\theta_1) = \hat{m}_{33}(\theta_1) = 0$ and the principal consults the references about 1 and 3 from 2. Agent 2 is lying about 1 and 3 in a way that the worse agent receives the better reference. Then $\pi_g^{so}(\hat{m}(\theta)) = 0$.

5.2 Partially honest agents

In the previous example, full implementation with mechanism π^{so} fails because agent 2 is indifferent between all her messages when she knows that $\pi_2^{so}(m_2, m_{-2}) = 0$ for all $m_2 \in M_2(\theta_2)$. In this case, the principal cannot infer the truth through references about the best applicants in every equilibrium because agent 2 has multiple best responses. If agents follow a tie-breaking rule when multiple messages maximize their expected probability of winning, then the principal can account for the tie breaking rule in her mechanism and deduce the truth.

One such tie-breaking rule is partial honesty as first introduced to the implementation literature by Dutta and Sen (2012). Partial honesty means that agents only lie if lying increases their expected probability of winning and otherwise they tell the truth. Such preferences have a lexicographic character: agents first care about maximizing their expected probability of winning, and second about telling the truth.

The assumption of partial honesty is justified in situations when agents first care about their own success and second that the chosen outcome matches the true state. For example, an employee could prioritize being the one promoted, but if he knows that he will not be promoted, then he is in favor of promoting the most qualified person. In academic environments, a researcher might care most about her own work being recognized, but if she knows that her work will not be selected, then she wants the best work to get recognition.

With partially honest agents, Maskin-/Bayesian-monotonicity and the revelation principle still fail in our setting. Dutta and Sen (2012) find that partial honesty can recover full implementation for complete information setting among the agents, if Maskin-monotonicity is not satisfied. Korpela (2014) shows that for a setting like ours without the hard evidence structure, incentive-compatibility of the principal's objective is sufficient and necessary for a fully implementing mechanism to exist if all agents are partially honest. The principal's objective in our setting is not incentive-compatible, as follows from the failure of the revelation principle. This implies that if there was no hard evidence in our setting, full implementation would be impossible even if all agents are partially honest.

In Definition 5, we present a mechanism which fully implements the principal's objective for class of graphs in our setup, if all agents are partially honest. Thus, our hard evidence structure recovers full implementation. It has not been determined which exact properties of our hard evidence structure are responsible for this and we leave this for future research.

Let Γ^h denote the game in which all agents are partially honest and everything else is as in the base game introduced in Section 2. With partially honest agents, the equilibrium definition has to be extended. The strategy profile $\hat{m}$ is an equilibrium of $\Gamma^h(\pi)$ if for all i, all θ_i and all $m_i \in M_i(\theta_i)$

1. $\Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i) \geq \Pi(m_i, \hat{m}_{-i}|\theta_i)$, and

2. $\hat{m}_i(\theta_i) = \theta_i$ if $\Pi(\theta_i, \hat{m}_{-i}|\theta_i) \geq \Pi(m_i, \hat{m}_{-i}|\theta_i)$.

In words, a strategy profile is an equilibrium, if given the others' strategies, every agent for each of her types chooses a message which maximizes her expected probability of winning, and which is the truth if the truth maximizes her expected probability of winning.

Before introducing mechanism π^{soh}, a tweaked version of π^{so} which yields full implementation for a class of graphs, we show with Example 4 that the previous mechanism π^{so} is not successful with partially honest agents because existence of an equilibrium is not guaranteed.

Example 4. Assume $n = 3$, L complete, and $b = .2$. Consider θ such that $d_1 = .4, d_2 = .5$ and $d_3 = .6$. Suppose $\hat{m}$ is an equilibrium of $\Gamma^h(\pi^{so})$. Then $\pi_1^{so}(\hat{m}(\theta)) = 1$, otherwise 1 would deviate to m_1' with $m_{11}' = .2$ such that $\pi_1^{so}(m_1', \hat{m}_{-1}(\theta_{-1})) = 1$. Hence, 2 and 3 say the truth because $\pi_2^{so}(\hat{m}(\theta)) = \pi_3^{so}(\hat{m}(\theta)) = 0$. If 2 and 3 say the truth, then $\pi_1^{so}(\theta_1, \hat{m}_{-1}(\theta_{-1})) = 1$ and 1 must say the truth as well. If all agents say the truth, however, 2 deviates to m_2' with $m_{22}' = .3$ such that $\pi_2^{so}(m_2', \hat{m}_{-2}(\theta_{-2})) = 1$. Thus no equilibrium exists for the complete network.

We tweak π^{so} to ensure equilibrium existence. The adjusted mechanism π^{soh} forms $B_1(m)$ and $B_2(m)$, in the same way as π^{so}. Mechanism π^{soh} includes another subset $B_3(m) \subseteq B_2(m)$ with the purpose of incentivizing the global minimum to "prove" that she is better than every neighbor or to conflict with some neighbor. We say agent i proves better than her neighbor j with message m_i if $m_{ij} - m_{ii} > 2b$ because in this case the principal knows with probability 1 that $d_i < d_j$. We say agent i conflicts with neighbor j if $m_{ij} \neq m_{jj}$ or $m_{ii} \neq m_{ji}$.

Definition 5. ***Applications-References-Honesty Mechanism*** π^{soh}

For any $m \in M$,

first choose $B_1(m) \subseteq N$ *such that* $i \in B_1(m)$ *if and only if* $m_{ii} = \min_{k \in N} m_{kk}$.

Second choose $B_2(m) \subseteq B_1(m)$ *such that*
if $\min_{k \in N} m_{kk} > 0$, *then* $B_2(m) = B_1(m)$, *and*
if $\min_{k \in N} m_{kk} = 0$, *then* $i \in B_2(m)$ *if and only if* $\bar{r}_i(m) = \min_{k \in B_1(m)} \bar{r}_k(m)$.

Third choose $B_3(m) \subseteq B_2(m)$ *such that* $i \in B_3(m)$ *if and only if 1) or 2) is satisfied:*
1) $m_{ii} \neq m_{ji}$ *or* $m_{ij} \neq m_{jj}$ *for some* $j \in N_i$
2) $m_{ij} - m_{ii} > 2b$ *for all* $j \in N_i$

Finally, choose $\pi^{soh}(m)$ *as follows.*

If $B_3(m) \neq \emptyset$*, then* $\pi_i^{soh}(m) = \frac{1}{|B_3(m)|}$ *for all* $i \in B_3(m)$*.*
The following allocations are "punishment allocations":
If $B_3(m) = \emptyset$ *and*

- $|B_2(m)| > 1$*, then* $\pi_i^{soh}(m) = \frac{1}{|B_2(m)|}$ *for all* $i \in B_2(m)$*.*
- $B_2(m) = \{i\}$*, then* $\pi_i^{soh}(m) = 0$*, and*
  - *if there is* $j \notin N_i$*, then* $\pi_j^{soh}(m) = \frac{1}{|N|-|N_i|-1}$ *for all* $j \notin N_i$*.*
  - *if all* $j \neq i$ *are in* N_i *and* $m_{ij} - m_{ii} > 2b$ *for some* $j \in N_i$*, then* $J = \{j \in N_i | m_{ij} - m_{ii} > 2b\}$ *and* $\pi_j^{soh}(m) = \frac{1}{|\,J\,|}$ *for all* $j \in J$*.*
  - *if all* $j \neq i$ *are in* N_i *and* $m_{ij} - m_{ii} \leq 2b$ *for all* $j \in N_i$*, then* $\pi_j^{soh}(m) = \frac{1}{|N|-1}$ *for all* $j \neq i$*.*

For any m, the construction of $B_1(m)$ and $B_2(m)$ for π^{soh} is the same as for π^{so}. For $B_3(m)$, the principal selects all agents from $B_2(m)$ who prove better than all their neighbors, or who conflict with some neighbor. For every equilibrium of Γ^h which we establish in the following, $B_3(m)$ is never empty and only agents who are in $B_3(m)$ are selected with positive probability. The outcomes for $B_3(m) = \emptyset$ are punishment allocations to prevent deviations from equilibria in which $B_3(m) \neq \emptyset$.

Our first result for π^{soh} is that π^{soh} partially implements the principal's objective for every L. This means that an equilibrium of $\Gamma^h(\pi^{soh})$ exists for all L. The extension of π^{so} by $B_3(m)$ successfully restores equilibrium existence when all agents are partially honest.

Proposition 5.3. *Mechanism* π^{soh} *partially implements the principal's objective in every* L*.*

The proof of proposition 5.3 consists of three steps. First, we define a strategy profile $\hat{m}^h$. Second, we show that $\pi_g^{soh}(\hat{m}^h(\theta)) = 1$ for all θ and third, that $\hat{m}^h$ is an equilibrium of $\Gamma^h(\pi^{soh})$.

We will refer to agent i with $d_i < d_j$ for all $j \in N_i$ and $|N_i| < n-1$ as a *local minimum with partial information*, to agent i with $d_i < d_j$ for all $j \in N_i$ and $|N_i| = n-1$ as a *local minimum with full information*, and to agent i with $d_i > d_j$ for some $j \in N_i$ as *non-minimal*. Observe that, given a type realization, if there exists a local minimum with full information, then this is the global minimum and all other agents are non-minimal. If there does not exist a local minimum with full information, then there exists at least one local minimum with partial information and some local minimum with partial information is the global minimum. Every local minimum with partial information expects with positive probability to be the global minimum. A neighbor of a local minimum is non-minimal.

Strategy profile $\hat{m}^h$ is such that every local minimum either proves that she is better than each of her neighbors or lies to the full extent about herself and her neighbors. Every non-minimal agent says the truth. Specifically, let $\hat{m}^h$ with $\hat{m}_i$ for all i be such that

- if i is non-minimal, then $\hat{m}_i(\theta_i) = \theta_i$.
- if i is a local minimum with partial information, then
  - $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$, and
  - $\hat{m}_{ij}(\theta_i) = \min\{d_j + b, 1\}$ for all $j \in N_i$,
    if $d_j - d_i \leq 2b$ for some $j \in N_i$, and
  - $\hat{m}_{ij}(\theta_i) = d_j$ for all $j \in N_i$, if $d_j - d_i > 2b$ for all $j \in N_i$.
- if agent i is a local minimum with full information, then
  - $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ and $\hat{m}_{ij}(\theta_i) = \min\{d_j + b, 1\}$ for all $j \in N_i$,
    if $d_j - d_i \leq 2b$ for some $j \in N_i$, and
  - $\hat{m}_i(\theta_i) = \theta_i$, if $d_j - d_i > 2b$ for all $j \in N_i$.

Lemma 5.4. *$\pi_g^{soh}(\hat{m}^h(\theta)) = 1$ for all θ.*

The formal proof is in the appendix, and we explain the intuition here.

Observe first that the global minimum is always a local minimum. Every neighbor of a local minimum is non-minimal and says the truth.

If $d_g > b$, g sends the unique best application. In case $d_j - d_g \leq 2b$ for some $j \in N_g$, g conflicts with a neighbor because she exaggerates. In case $d_j - d_g > 2b$ for all $j \in N_g$, g proves better than all her neighbors. Thus $B_3(\hat{m}^h(\theta)) = \{g\}$.

If $d_g \leq b$, every local minimum i with $d_i \leq b$ sends a best application. The global minimum uniquely receives the min-max reference because references about all local minima are from truthful non-minimal agents. Again, g conflicts with a neighbor if $d_j - d_g \leq 2b$ for some $j \in N_g$ because g exaggerates, and proves better than all neighbors if $d_j - d_g > 2b$ for all $j \in N_g$. Thus $B_3(\hat{m}^h(\theta)) = \{g\}$.

Lemma 5.5. *The strategy profile $\hat{m}^h$ is an equilibrium of $\Gamma^h(\pi^{soh})$.*

The formal proof is provided in the appendix. We explain the intuition here. We show that given $\hat{m}_{-i}$, $\hat{m}_i(\theta_i)$ maximizes i's expected probability of winning and that the truth does not maximize i'expected probability of winning if $\hat{m}_i(\theta_i)$ is not the truth for every θ_i. From this we can conclude that $\hat{m}^h$ is an equilibrium of $\Gamma^h(\pi^{soh})$.

Agent i who is a local minimum with full information knows that she is the global minimum and that $\pi_i^{soh}(\hat{m}^h(\theta)) = 1$. Trivially, her message maximizes her expected probability of winning. Agent i only does not say the truth, if $d_j - d_i \leq 2b$ for some $j \in N_i$. If she then deviates to the truth m_i', i neither conflict with a neighbor nor does she prove better than all of her neighbors any longer. Thus $B_2(m_i', \hat{m}_{-i}(\theta_{-i})) = \{i\}$ and $B_3(m_i', \hat{m}_{-i}(\theta_{-i})) = \emptyset$ such that $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) = 0$.

Agent i who is a local minimum with partial information expects with positive probability that she is g and that $\pi_i^{soh}(\hat{m}^h(\theta)) = 1$. We show in the appendix that, in case i is not g, $B_3(m_i', \hat{m}_{-i}(\theta_{-i})) = \{g\}$ and $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) =$

0 for any $m_i' \in M_i(\theta_i)$. In each case $\hat{m}_i(\theta_i)$ maximizes her probability of winning and thus it maximizes her expected probability of winning.

If $d_i > 0$, then i expects with positive probability that she is g and that there is another local minimum j with $d_j < d_i + b$. If this is the case and i deviates to the truth m_i', then i loses against j and $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) = 0$. If $d_i = 0$, then i knows that she is g and that $\pi_i^{soh}(\hat{m}^h(\theta)) = 1$. Agent i only does not say the truth, if $d_j - d_i \leq 2b$ for some $j \in N_i$. If she then deviates to the truth m_i', i neither conflict with a neighbor nor does she prove better than all of her neighbors. Thus $B_2(m_i', \hat{m}_{-i}(\theta_{-i})) = \{i\}$ but $B_3(m_i', \hat{m}_{-i}(\theta_{-i})) = \emptyset$ such that $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) = 0$. Hence, if i does not say the truth, then deviating to the truth strictly decreases her expected probability of winning.

Agent i who is non-minimal knows that she is not g and that $\pi_i^{soh}(\hat{m}^h(\theta)) = 0$. There is no $m_i' \in M_i(\theta_i)$ such that $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) > 0$ and thus the truth maximizes i's expected probability of winning.

If $d_i > b$, then any application of i is worse than m_{gg}. If $d_i \leq b$ and g is a neighbor of i, then $\bar{r}_i(m_i', \hat{m}_{-i}(\theta_{-i})) = d_i + b > d_g + b$, and in case g is not a neighbor of i, then $\bar{r}_i(m_i', \hat{m}_{-i}(\theta_{-i})) \geq d_i > \bar{r}_g(m_i', \hat{m}_{-i}(\theta_{-i})) = d_g$ for all $m_i' \in M_i(\theta_i)$. Thus, $i \notin B_3(m_i', \hat{m}_{-i}(\theta_{-i}))$ for any $m_i' \in M_i(\theta_i)$.

Agent i can also not achieve a positive probability of winning through punishment allocations for any $m_i' \in M_i(\theta_i)$. For agent i to cause $B_3(m_i', \hat{m}_{-i}(\theta_{-i})) = \emptyset$ and $B_2(m_i', \hat{m}_{-i}(\theta_{-i})) = \{j\}$ with $j \neq i$, it is necessary that $j \in N_i$ and $d_i - d_j \leq 2b$. If there is $k \notin N_j$ or $d_k - d_j > 2b$ for $k \in N_j$, then $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) = 0$. If there is no $k \notin N_j$ or $d_k - d_j \leq 2b$ for all $k \in N_j$, then j conflicts with all of her neighbors, and there is no deviation by i such that $B_3(m_i', \hat{m}_{-i}(\theta_{-i})) = \emptyset$ and $B_2(m_i', \hat{m}_{-i}(\theta_{-i})) = \{j\}$.

Hence, $\hat{m}^h$ is an equilibrium of $\Gamma^h(\pi^{soh})$.

Next, we find that π^{soh} fully implements the principal's objective for a class of graphs. We provide sufficient conditions for the properties of L for π^{soh} to be fully implementing (Theorem 5.6).

Theorem 5.6. *Mechanism π^{soh} fully implements with partially honest agents in L, if L is connected and such that for all $i \in N$ and all $j \in N_i$*

1. $N_j \backslash i \subset N_i$*, or*

2. j *is linked to every* $k \neq j$*.*

In words, the sufficient conditions on L for π^{soh} to be fully implementing with partially honest agent are the following: First, L is connected, and, second, for every agent i and every neighbor $j \in N_i$, it is true that i is linked to all the other neighbors of j or that j is linked to every other agent.

Let superscript $*$ denote that L satisfies the conditions of Theorem 5.6. We first point out two important implications regarding the information structure in L^*. If agent i is linked to all the other neighbors of j, then i knows what j knows and thus i knows θ_j. If agent i is not linked to all the other neighbors of j, then i does not know θ_j; agent j, however, has full information and knows every agents' type.

The formal proof for Theorem 5.6 is in the appendix and we provide the intuition here. The first step in the proof is to show that there is at least one agent who has full information in any L^*. Second, we prove that every equilibrium $\hat{m}$ of $\Gamma^h(\pi^{soh})$ is such that with probability 1, $B_3(\hat{m}(\theta)) \neq \emptyset$ and thus for all $i \in N$, $\pi_i^{soh}(\hat{m}(\theta)) > 0$ if and only if $i \in B_3(\hat{m}(\theta))$. The third part of the proof then shows that every equilibrium $\hat{m}$ of $\Gamma^h(\pi^{soh})$ is such that $B_3(\hat{m}(\theta)) = \{g\}$ with probability 1.

This third part consists of several sub-considerations.

If g is agent i with full information and $\pi_i^{soh}(\hat{m}(\theta)) < 1$, then there is a deviation m_i' such that $\pi_i^{soh}(m_i', \hat{m}_{-i}(\theta_{-i})) = 1$.

Next, we consider that g is agent i with partial information. If $d_i > b$ and $\pi_i^{soh}(\hat{m}(\theta)) < 1$ with positive probability, then there is a deviation m_i' such that $\Pi(m_i', \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$. Finally we suppose $d_i \leq b$. We find that in this case with probability 1 every equilibrium $\hat{m}$ has the following properties: Every local minimum l with $d_l \leq b$ sends $\hat{m}_{ll}(\theta_l) = 0$ and $\hat{m}_{lj}(\theta_l) > d_l + b$ for all $j \in N_l$ with $d_j \leq b$. Every j who is a neighbor of a local minimum l with $d_l \leq b$, or who is non-minimal with $d_j \leq b$ says the truth because $\Pi(\hat{m}_j(\theta_j), \hat{m}_{-j}|\theta_j) = 0$. Thus g is local minimum l with $d_l \leq b$ who sends $\hat{m}_{ll}(\theta_l) = 0$, has $\bar{r}_l(\hat{m}(\theta)) < \bar{r}_k(\hat{m}(\theta))$ for all $k \neq l$ in $B_1(\hat{m}(\theta))$ and hence $B_2(\hat{m}(\theta)) = \{g\}$. From $B_3(\hat{m}(\theta)) \neq \emptyset$ and $B_3(\hat{m}(\theta)) \subseteq B_2(\hat{m}(\theta))$

follows that $B_2(\hat{m}(\theta)) = \{g\}$. This concludes the proof.

Corollary 5.7 provides further insight into which graphs satisfy the conditions of theorem 5.6.

Corollary 5.7. *L satisfies the conditions of Theorem* 5.6 *if and only if given L, there is a partition of N into cliques $C_1, ..., C_K$ with $K \geq 1$ such that*

- *every $i \in C_1$ is linked to all $j \neq i$.*
- *every $i \in C_k$ is linked to $j \in C_l$ if and only if $k = l$ for all $k, l > 1$.*

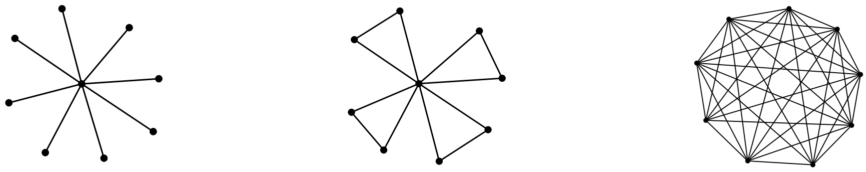
In words, every agent in C_1 is linked to every other $j \in N$, and every agent in C_k is linked to every other agent in C_k and to every other agent in C_1, and is not linked to any agent in other cliques. The proof is straightforward.

Proof. Take any L such that the conditions of Theorem 5.6 are satisfied. First, assign every i who is linked to all $j \neq i$ to C_1. Second, assign $i \in N \backslash C_1$ and every $j \in N_i$ who is not linked to all $k \neq j$ to C_2; observe that all agents in C_2 are fully linked among each other because $N_j \backslash i \subset N_i$ and $N_i \backslash j \subset N_j$ for all $j \in C_2$ as no agent in C_2 is linked to all other agents in N; moreover agents in C_2 do not have any links to agents in $\{N \backslash C_1\} \backslash C_2$. Third, assign $i \in \{N \backslash C_1\} \backslash C_2$ and every $j \in N_i$ who is not linked to all $k \neq j$ to C_3 and so on, until all agents are assigned. This is a partition as described by Corollary 5.7.

Consider L such that there is a partition of N which satisfies the conditions of Corollary 5.7. Graph L is connected because of C_1. Each $i \in C_1$ is linked to all the neighbors of each of her neighbors because i is linked to every other agent. Each $i \in C_k$ for all $k > 1$ is linked to all the other neighbors of $j \in N_i$ if $j \in C_k$, or j has full information if $j \in C_1$; this is true for all $j \in N_i$. □

The following graphs are examples of L^*. Note that any L^* has maximum diameter 2.[5]

[5]Suppose agent i is linked to one agent k who has full information. This means k is linked to every agent and thus i has at most distance 2 to every other agent. Suppose



(a) Star (b) Dutch Windmill (c) Complete Graph

Figure 2: Examples of L^* for $n = 9$

The mechanism π^{soh}, however, does not fully implement for all graphs. In particular, the full implementation property of mechanism π^{soh} is non-monotonic in the number of links of the graphs. For an incomplete supergraph of the star, full implementation with π^{soh} fails. For the complete graph, full implementation with π^{soh} is recovered. This shows that more information together with more communication is not always beneficial for full implementation.

Proposition 5.8. *There is L which is a supergraph of the star and a subgraph of the complete graph such that π^{soh} does not fully implement in L.*

For the proof, we first define L and n. Second, we present a strategy profile $\hat{m}^l$ and show that $\hat{m}^l$ is an equilibrium of $\Gamma^h(\pi^{soh})$. Finally, we find $\Theta' \subset \Theta$ with positive measure such that $\pi_g^{soh}(\hat{m}^l(\theta)) = .5$ for all $\theta \in \Theta'$.

Proof. Let $n = 5$ and L be as in Figure 3.

Define the following strategy profile $\hat{m}^l$.

Let $\hat{m}_5(\theta_5)$ be such that

- if $d_5 < d_j$ for all $j \neq 5$ and $d_j - d_5 \leq 2b$ for some $j \neq 5$, then
  $\hat{m}_{55}(\theta_5) = \max\{0, d_5 - b\}$ and $\hat{m}_{5j}(\theta_5) = \min\{d_j + b, 1\}$ for all $j \neq 5$,
- if $d_5 < d_j$ and $d_j - d_5 > 2b$ for all $j \neq 5$, or if $d_5 \geq d_j$ for some $j \neq 5$, then $\hat{m}_5(\theta_5) = \theta_5$.

agent i is not linked to an agent with full information. This means that agent i is linked to all the neighbors of each of her neighbors. Then agent i is linked to every agent and has distance 1 to every other agent because L is connected.

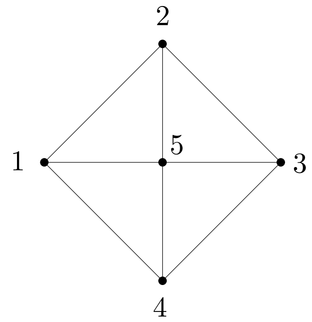


Figure 3

In words, agent 5 fully exaggerates, if 5 is a local minimum and some neighbor's distance is relatively close to hers. Agent 5 says the truth, if she is a local minimum and all neighbors' distances are far, or if she is non-minimal.

Let $\hat{m}_i(\theta_i)$ for all $i \neq 5$ be such that

- if $d_i < d_j$ for all $j \in N_i$ and
  - if $d_j - d_i \leq 2b$ for some $j \in N_i$, then $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ and $\hat{m}_{ij}(\theta_i) = \min\{d_j + b, 1\}$ for all $j \in N_i$,
  - if $d_j - d_i > 2b$ for all $j \in N_i$, then $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ and $\hat{m}_{ij}(\theta_i) = d_j$ for all $j \in N_i$.
- if $d_i \geq d_j$ for some $j \in N_i$ and
  - if $d_i > b$, or $d_j = 0$ for some $j \in N_i$, or $d_j > b$ for all $j \in N_i \backslash 5$, then $\hat{m}_i(\theta_i) = \theta_i$,
  - if $d_i \leq b$ and $d_j > 0$ for all $j \in N_i$ and $d_j \leq b$ for some $j \in N_i \backslash 5$, then $\hat{m}_{ii}(\theta_i) = 0$, $\hat{m}_{ij}(\theta_i) = b$ for all $j \in N_i$ with $d_j \leq b$, and $\hat{m}_{ij}(\theta_i) = d_j$ for all $j \in N_i$ with $d_j > b$.

In words, agent $i \neq 5$ who is a local minimum sends the best possible application, and if some neighbor's distance is close, worst possible references, or if each neighbor's distance is far, truthful references. If agent $i \neq 5$ is non-minimal, and is sufficiently bad ($d_i > b$), or knows to be linked to g ($d_j = 0$ for some $j \in N_i$), or all $j \in N_i \backslash 5$ are sufficiently bad ($d_j > b$), then i says the truth. If agent $i \neq 5$ is non-minimal, is sufficiently good ($d_i \leq b$), does not know to be linked to g ($d_j > 0$ for all $j \in N_i$), and some $j \in N_i \backslash 5$ is sufficiently good ($d_j \leq b$), then i sends the best application equal to 0, false references equal to b about sufficiently good neighbors, and true references about sufficiently bad neighbors.

Next, we show that $\hat{m}^l$ is an equilibrium of $\Gamma^h(\pi^{soh})$. A more detailed version of this is in the appendix.

Let $\hat{m}^l(\theta) = m$.

First, consider the strategy of agent 5. If 5 is g, then $\pi_5^{soh}(m) = 1$. If 5 is not g, then $\pi_5^{soh}(m) = 0$, and for no m_5', $5 \in B_2(m_5', m_{-5})$ because either $m_{gg} < d_5 - b$ or $m_{gg} = 0$ and $m_{g5} > d_g + b$. Agent 5 cannot gain via punishment allocations neither because if $B_2(m_5', m_{-5}) = \{j\}$ for $j \neq 5$ and $B_3(m_5', m_{-5}) = \emptyset$, then $\pi_k^{soh}(m_5', m_{-5}) = 1$ for $k \notin N_j$. Agent 5 only does not say the truth, if 5 is g and $d_j - d_5 \leq 2b$ for some $j \neq 5$. Deviating to the truth then either leads to $B_2(m_5', m_{-5}) = \{5\}$ and $B_3(m_5', m_{-5}) = \emptyset$, or $5 \notin B_3(m_5', m_{-5}) \neq \emptyset$. In either case, $\pi_5^{soh}(m_5', m_{-5}) = 0$.

Second, consider the strategy of agent $i \neq 5$, if she is a local minimum. Observe that no $j \in N_i$ is in $B_2(m)$ because either $m_{ii} < m_{jj}$ or both $m_{ii} = 0$ and $m_{ij} > d_i + b$.

Suppose i is g. If $j \notin N_i$ says the truth, then $B_3(m) = \{i\}$ and $\pi_i^{soh}(m) = 1$. If $j \notin N_i$ does not say the truth, then $m_{ii} = m_{jj} = 0$. If all $j \in N_i$ send truthful references, then $\bar{r}_i(m) < \bar{r}_j(m)$ for $j \notin N_i$ and hence $B_3(m) = \{i\}$. If some $j \in N_i$ sends references equal to b, then $\bar{r}_i(m) = \bar{r}_j(m)$ for $j \notin N_i$ and $B_3(m) = \{i, j\}$ for $j \notin N_i$. In the last case, $j \in B_3(m_i', m_{-i})$ for $j \notin N_i$ and $\pi_i^{soh}(m_i', m_{-i}) \leq \pi_i^{soh}(m)$ for all $m_i' \in M_i(\theta_i)$.

Suppose i is not g. Then $j \notin N_i$ is g. If $d_i > b$, or if $d_i \leq b$ and all $j \in N_i$

say the truth, then $B_3(m) = \{g\}$ and $\pi_i(m) = 0$ and there does not exist $m_i' \neq m_i$ such that $B_3(m_i', m_{-i}) \neq \{g\}$. If $d_i \leq b$ and some $j \in N_i$ does not say the truth, then $\bar{r}_i(m) = \bar{r}_g(m) = b$ and $B_3(m) = \{g, i\}$ and there does not exist $m_i' \neq m_i$ such that $g \notin B_3(m_i', m_{-i})$.

Thus m_i maximizes i's expected probability of winning. We show next that deviating to the truth, if m_i is not the truth, strictly decreases i's expected probability of winning.

If $d_i > 0$, then i expects with positive probability to be g and that $j \notin N_i$ is also a local minimum with $d_j \in (d_i, \min\{d_i + b, 1\})$ such that $\pi_i^{soh}(m) \geq \frac{1}{2}$. A deviation to the truth m_i' causes i to lose against $j \notin N_i$ and $\pi_i^{soh}(m_i', m_{-i}) = 0$. If $d_i = 0$, i is g and $\pi_i^{soh}(m) = 1$. In this case, agent i only does not say the truth in case $d_j \leq 2b$ for some $j \in N_i$. A deviation to the truth m_i' leads to $i \in B_2(m_i', m_{-i})$ and $B_3(m_i', m_{-i}) = \emptyset$ such that $\pi_i^{soh}(m_i', m_{-i}) = 0$.

Third, consider the strategy of agent $i \neq 5$, if she is non-minimal and says the truth such that $\pi_i^{soh}(m) = 0$. In this case, there is no $m_i' \in M_i(\theta_i)$ such that $i \in B_2(m_i', m_{-i})$ because $m_{gg} < m_{ii}'$ or $\bar{r}_g(m_i', m_{-i}) < \bar{r}_i(m_i', m_{-i})$ for all $m_i' \in M_i(\theta_i)$. Agent i can also not gain through the punishment allocations. For all $m_i' \in M_i(\theta_i)$ and for $j \notin N_i$, $B_3(m_i', m_{-i}) = \{j\}$ if j is g, or $j \notin B_2(m_i', m_{-i})$ if j is not g, and $B_3(m_i', m_{-i}) = \{5\}$ if 5 is g, or $5 \notin B_1(m_i', m_{-i})$ if 5 is not g. So the truth maximizes agent i's expected probability of winning.

Finally, consider the strategy of agent $i \neq 5$, if she is non-minimal and does not say the truth. Suppose $g \notin N_i$. If $d_g > 0$, then $\bar{r}_i(m) = \bar{r}_g(m) = b$ such that $B_3(m) = \{g, i\}$. For any $m_i' \neq m_i$, $g \in B_3(m_i', m_{-i})$. If $d_g = 0$, then $B_3(m) = \{g\}$ and $B_3(m_i', m_{-i}) = \{g\}$ for all $m_i' \neq m_i$.

Suppose $g \in N_i$. Then $i \notin B_3(m) \neq \emptyset$ and $\pi_i(m) = 0$. For all $m_i' \neq m_i$, $k \notin B_2(m_i', m_{-i})$ because $m_{gg} = 0$ and $m_{gk} > d_g + b$ for all $k \in N_g$. If 5 is g, then $B_3(m_i', m_{-i}) = \{g\}$ for all $m_i' \neq m_i$ because 5 conflicts with all $j \in N_5 \backslash i$.

Thus m_i maximizes agent i's expected utility. A deviation to the truth strictly decreases agent i's expected utility because i assigns strictly positive probability to $g \notin N_i$ and $d_g > 0$ such that $B_3(m) = \{g, i\}$. For this case, deviating to the truth m_i' results into $B_3(m_i', m_{-i}) = \{g\}$.

For all θ, either $B_3(\hat{m}^l(\theta)) = \{g\}$ or $B_3(\hat{m}^l(\theta)) = \{g, j\}$ where $j \notin N_g$ and thus $\pi_g^{soh}(\hat{m}^l(\theta)) = 1$ or $\pi_g^{soh}(\hat{m}^l(\theta)) = \frac{1}{2}$.

If $0 < d_1 < d_j < d_5 \leq b$ for all $j = 2, 3, 4$, then $\hat{m}_5(\theta_5) = \theta_5$, $\hat{m}_{ii}(\theta_i) = 0$ for all $i \neq 5$, $\hat{m}_{1j}(\theta_1) = d_j + b$ for all $j \in N_1$, and $\hat{m}_{ij}(\theta_1) = b$ for all $j \in N_i$ and all $i = 2, 3, 4$. Thus $B_3(\hat{m}^l(\theta)) = \{1, 3\}$. The lie is optimal for 2 and 4 because each of them assigns a positive probability to the other one being g which would mean $B_3(\hat{m}^l(\theta)) = \{2, 4\}$. $\square$

5.3 Noisy communication

When communication is noisy, then the simpler mechanism π^{so} achieves full implementation in expectation for every L, independent of partial honesty. The reason is that fully exaggerating about oneself and about neighbors becomes the unique expected utility maximizing message for any strategy profile of others. For simplicity, we drop the assumption of partial honesty again for this section.

We define noisy communication in the following way. For any m_i which agent i sends the principal receives message $\tilde{m}_i$ with $\tilde{m}_{ik} = m_{ik} + \epsilon_{ik}$ for all k and all i where ϵ_{ik} is a statement-specific noise term which is independently and identically drawn from a continuous distribution with mean 0 and full support on $[-m_{ik}, 1 - m_{ik}]$ with likelihood strictly decreasing in the absolute size of the error. We denote the game with noisy communication and everything else as in the base game by Γ^n. Thus, an agent's action set in $\Gamma^n(\pi)$ is still restricted by the limit to lying. However, for any statement m_{ik} chosen by agent i, the principal receives the noisy version $\tilde{m}_{ik}$ where $\tilde{m}_{ik}$ can take any value in $[0, 1]$ with higher likelihood for values closer to m_{ik}. Let $\tilde{m}$ denote the message profile received from the principal when agents choose message profile m. The outcome then is a function of $\tilde{m}$ and agent i's probability of receiving the prize is given by $\pi_i(\tilde{m})$.

Proposition 5.9. *The unique equilibrium of $\Gamma^n(\pi^{so})$ for any L is $\hat{m}$ such* *that* $\hat{m}_{ii}(\theta_i) = \max\{0, d_i - b\}$ *and* $\hat{m}_{ij}(\theta_i) = \min\{d_j + b, 1\}$ *for all* $j \in N_i$*, all* θ_i *and all* i*. The mechanism* π^{so} *fully implements the principal's objective in expectation for all* L*.*

The proof follows immediately from the following observation. Any agent i for any θ_i, any $m_i \in M_i(\theta_i)$ and any $\hat{m}_{-i}$ expects with positive probability that $\tilde{m}_{ii} \leq \tilde{m}_{kk}$ for all $k \neq i$ and that if $\tilde{m}_{ii} = 0$, then $\bar{r}_i(\tilde{m}) < \bar{r}_k(\tilde{m})$ for all $k \neq i$ for whom $\tilde{m}_{kk} = 0$. This probability is maximized if and only if $m_{ii} = \max\{0, d_i - b\}$ and $m_{ij} = \min\{d_j + b, 1\}$ for all $j \in N_i$. This means given π^{so}, the unique expected utility maximizing message for any $\hat{m}_{-i}$ is $m_{ii} = \max\{0, d_i - b\}$ and $m_{ij} = \min\{d_j + b, 1\}$ for all $j \in N_i$.

Given the unique equilibrium $\hat{m}$ of $\Gamma^n(\pi^{so})$, the expected message profile which the principal receives is $E[\tilde{m}] = \hat{m}(\theta)$ and thus the expected probability with which the global minimum gets the prize for any θ is $E[\pi_g^{so}(\tilde{m})] = \pi_g^{so}(E[\tilde{m}]) = \pi_g^{so}(\hat{m}(\theta)) = 1$.

6 Conclusion

This paper studies how a principal can allocate a prize to the best agent in every equilibrium, if transfers and disposal of the prize are not possible, and if agents have a limit to lying. The limit to lying is essential – without it, full implementation would be impossible in our setting. Two intuitive mechanisms which allocate the prize as a function of either applications or references only do not achieve full implementation. We propose mechanism π^{so} which allocates the prize as a function of the best application and worst references. This mechanism partially implements the principal's objective in all networks in dominant strategies and fully implements in the complete network. If agents are partially honest, then mechanism π^{soh} is partially implementing in all networks, and fully implementing in for a class of networks. If communication is noisy, then π^{so} fully implements in expectation in all networks.

We achieve all our implementation results via untruthful equilibria. This raises an important issue. If agents have a limit to lying, then the concern in the literature about truthful revelation might not be justified. The principal can equally well, maybe better, achieve her goals, if agents are dishonest. Dishonesty is not harmful in our setting because agents are dishonest in predictable ways. We can find many real world examples where agents are dishonest in predictable ways. If an individual is asked to submit supporting materials for an application, then anyone most likely submits the most positive evidence, although this evidence is an outlier to the top – think of your best teaching evaluations versus your average teaching evaluation. A concern for further research is then how the principal can ensure a limit to lying and determine its extent.

Another question raised by this paper is the effect of tie-breaking rules other than partial honesty in case agents have multiple best replies. For example, favoritism or spitefulness are two important behavioral considerations. Moreover, the assumption that knowledge between two linked agents is perfect should be challenged. If knowledge is imperfect, then the number of references for each individual should gain importance. A theoretical question which needs more investigation is which exact properties of hard evidence are important for full implementation in Bayesian settings.

Appendix

Proof of Lemma 5.4.

Proof. For any θ, g is a local minimum. Every local minimum sends her best possible application unless she has full information and $d_j - d_g > 2b$ for all $j \in N_g$ in which case she says the truth. This implies that if $\hat{m}_{gg}(\theta_g) > 0$, then $\hat{m}_{gg}(\theta_g)$ is the unique best application and $B_1(\hat{m}^h(\theta)) = B_2(\hat{m}^h(\theta)) = \{g\}$. If $\hat{m}_{gg}(\theta_g) = 0$, then $g \in B_1(\hat{m}^h(\theta))$ and g uniquely receives the min-max reference among all agents in $B_1(\hat{m}^h(\theta))$ because the neighbors of every local minimum are non-minimal and send truthful references about all agents in $B_1(\hat{m}^h(\theta))$. Thus $B_2(\hat{m}^h(\theta)) = \{g\}$ as well. It is left to show that g is also in $B_3(\hat{m}^h(\theta))$. The global minimum lies about her neighbors if $d_j - d_g \leq 2b$ for $j \in N_g$. Then g conflicts with her neighbors because her neighbors are non-minimal and send truthful applications. If $d_j - d_g > 2b$ for all $j \in N_g$, then the global minimum says the truth about her neighbors and $\hat{m}_{gj}(\theta_g) - \hat{m}_{gg}(\theta_g) > 2b$ for all $j \in N_g$. Hence, $B_3(\hat{m}^h(\theta)) = \{g\}$ and $\pi_g^{soh}(\hat{m}^h(\theta)) = 1$ for all θ. $\square$

Proof of Lemma 5.5.

Proof. Let $\hat{m}^h = \hat{m}$ and $\hat{m}(\theta) = m$. We show in turn that, given $\hat{m}_{-i}$, m_i maximizes i's expected probability of winning and $m_i = \theta_i$ if the truth maximizes i's probability of winning for $\theta_i =$*local minimum with full information*, $\theta_i =$*local minimum with partial information*, and $\theta_i =$*non-minimal*.

Consider first that agent i is a local minimum with full information. Then agent i knows that she is the global minimum and that $\pi_i(m) = 1$. Trivially, m_i maximizes i's expected probability of winning. If m_i is not the truth, then the distance of some neighbor $j \in N_i$ is close to d_i. By unilaterally deviating to the truth m_i', agent i would not conflict with any of her neighbors anymore because her neighbors are non-minimal and say the truth. At the same time, agent i would not prove to be better than all of her neighbors because $d_j - d_i \leq 2b$ for some $j \in N_i$ and hence $i \notin B_3(m_i', m_{-i})$ but $B_2(m_i', m_{-i}) = \{i\}$. Then agent i would be selected with probability 0 and deviating to the truth would strictly decrease i's expected probability of winning.

Consider second that agent i is a local minimum with partial information. Then agent i expects with positive probability to be the global minimum and hence that $\pi_i(m) = 1$. If she is not the global minimum, then $\pi_i(m) = 0$ and another local minimum agent g who is not her neighbor is the global minimum. For any other $m'_i \neq m_i$, first, g still sends the overall best application because m_{ii} is already i's best application, second, g still receives the min-max reference among agents in $B_1(m'_i, m_{-i})$ because i can neither change the references about herself nor about g or any other agent in $B_1(m)$ who are all local minima , and third, g still conflicts with every neighbor $j \in N_g$ for whom $d_j - d_g \leq 2b$. Hence for any $m'_i \neq m_i$, $B_3(m'_i, m_{-i}) = \{g\}$. Thus m_i maximizes i's expected probability of winning.

Next, observe that $\Pi(m_i, \hat{m}_{-i}|\theta_i) > \Pi(\theta_i, \hat{m}_{-i}|\theta_i)$ if m_i is not the truth. Agent i expects with strictly positive probability that she is the global minimum and that there exists another local minimum j with $d_j \in (d_i, d_i + b)$ who sends $m_{jj} = \max\{0, d_j - b\}$. Suppose this is indeed the case. If $d_i > 0$ and i deviates to the truth $m'_i = \theta_i$ with $m'_{ii} = d_i$, then her application is worse than j's and $B_3(m'_i, m_{-i}) \neq \emptyset$ 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 $m'_i = \theta_i$, agent i still sends the best application and receives the min-max reference but neither conflicts with a neighbor nor proves better than all her neighbors. Thus $B_3(m'_i, m_{-i}) = \emptyset$ and $B_2(m'_i, m_{-i}) = \{i\}$ such that $\pi_i(m'_i, m_{-i}) = 0$.

Consider third that agent i is non-minimal. Agent i knows that she is not the global minimum g and that $\pi_i(m) = 0$. We show that there does not exist a message m'_i such that $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > 0$.

First, observe that there does not exist m'_i such that $i \notin B_2(m'_i, m_{-i})$ and hence i never gets selected with $\pi_i(m'_i, m_{-i}) > 0$ as a member of $B_2(m'_i, m_{-i})$ or $B_3(m'_i, m_{-i})$ for any m'_i. If $d_i > b$, then agent i's best feasible application is strictly worse than the best application of the global minimum and hence

$i \notin B_1(m'_i, m_{-i})$ for any 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, $i \notin B_2(m'_i, m_{-i})$ for any m'_i: If g is a neighbor of i, then the maximum reference about i is $\overline{r}_i(m'_i, m_{-i}) = d_i + b$ and i can increase the maximum reference about g to at most $\overline{r}_g(m'_i, m_{-i}) = d_g + b$ which is still strictly less than $\overline{r}_i(m'_i, m_{-i})$. If g is not a neighbor of i, then i cannot increase the maximum reference about g and $\overline{r}_g(m'_i, m_{-i}) = d_g < \overline{r}_i(m'_i, m_{-i}) \in \{d_i, d_i + b\}$ for any m'_i.

Second, observe 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 $j \neq i$. For agent j to be in $B_2(m'_i, m_{-i})$ and not in $B_3(m'_i, m_{-i})$, j must send the best application, must receive the min-max reference among all agents in $B_1(m'_i, m_{-i})$ in case the best application is zero, and must not conflict with any neighbor where $m_{jk} - m_{jj} \leq 2b$ for some $k \in N_j$, given (m'_i, m_{-i}). The only candidate agents j to be in $B_2(m'_i, m_{-i})$ are those who did already sent $m_{jj} = \min_{k \in N} m_{kk}$ before any deviation of agent i. These are the global minimum g and other local minima l with $d_l \leq b$. Before any deviation of i, g and every l conflict with each of their neighbors $k \in N_j$ for whom $m_{jk} - m_{jj} \leq 2b$ for $j = l, g$. In order for $j \in B_2(m'_i, m_{-i})$ and $j \notin B_3(m'_i, m_{-i})$, m'_i must be such that 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, i must be a neighbor of 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, $\Pi(m'_i, \hat{m}_{-i}|\theta_i) = 0$ for every $m'_i \neq m_i$ and the true message m_i maximizes i's expected probability of winning.

Then $\hat{m}^h$ is an equilibrium. $\square$

Proof of Theorem 5.6.

Proof. Let the graph be connected and such that for all i and all $j \in N_i$, $N_j \backslash i \subset N_i$, or j is linked to all $k \neq j$. If i is linked to all neighbors $k \neq i$ of her neighbor $j \in N_i$, then i knows j's type. If i is not linked to all neighbors $k \neq i$ of $j \in N_i$ but j is linked to all $k \neq j$, then i does not know j's type. Agent j however knows every agent's type.

First, we show that there exists at least one agent who is linked to every other agent and thus has full information. If $N_j \backslash i \subset N_i$ for all $j \in N_i$, then i has full information. To see why, suppose i does not have full information. But then there must exist some $k \in N_j$ for some $j \in N_i$ to whom i is not linked, meaning $k \notin N_i$, because the graph is connected. This is a contradiction. If $N_j \backslash i \subset N_i$ not for all $j \in N_i$, then there is some $j \in N_i$ who has full information. Let F be the set of all agents with full information in L.

Second, we show that every equilibrium $\hat{m}$ is such that $B_3(\hat{m}(\theta)) \neq \emptyset$ with probability 1.

Suppose to the contrary that there is an equilibrium $\hat{m}$ such that $B_3(\hat{m}(\theta)) = \emptyset$ and $f \in B_2(\hat{m}(\theta))$ for some θ and $f \in F$. We show that there exist m'_f such that $\Pi(m'_f, \hat{m}_{-f}|\theta_f) > \Pi(\hat{m}_f(\theta_f), \hat{m}_{-f}|\theta_f)$ and thus $\hat{m}$ is not an equilibrium.

As f knows the type of every agent, f knows $m = \hat{m}(\theta)$ and that $B_3(m) = \emptyset$ and $f \in B_2(m)$ and thus $\pi_f(m) \leq \frac{1}{2}$. As $B_3(m) = \emptyset$ and $f \in B_2(m)$, m must be such that f sends the best application, gets the min-max reference if $m_{ff} = 0$, and does neither conflict with any neighbor nor proves better than all of her neighbors.

If $m_{ff} > \max\{0, d_f - b\}$, then f can deviate to $m'_{ff} = \max\{0, d_f - b\}$ and $m'_{fj} = m_{fj}$ for all $j \in N_f$. The deviation leads to $B_3(m'_f, m_{-f}) = \{f\}$ and $\pi_f(m'_f, m_{-f}) = 1$.

If $m_{ff} = \max\{0, d_f - b\}$, then f can deviate to $m'_{ff} = m_{ff}$ and $m'_{fk} > 0$ such that $m'_{fk} \neq m_{fk}$ for all $k \notin B_2(m)$ and $m'_{fj} = m_{fj}$ for all $j \neq k, f$. This leads to $B_3(m'_f, m_{-f}) = \{f\}$ and $\pi_f(m'_f, m_{-f}) = 1$. If every $k \neq f$ is also in $B_2(m)$, then deviating to $m'_{ff} = m_{ff}$ and $m'_{fk} > 0$ such that $m'_{fk} \neq m_{fk}$

for exactly one $k \neq f$ and $m'_{fj} = m_{fj}$ for all $j \neq k, f$. Then $\pi_f(m'_f, m_{-f}) \geq 1/2 > 1/n = \pi_f(m)$ because after f's deviation either $B_3(m'_f, m_{-f}) = \{f\}$ (if $m_{ff} = 0$) or $B_3(m'_f, m_{-f}) = \{f, k\}$ (if $m_{ff} > 0$).

Thus, there always exists a deviation m'_f such that $\Pi(m'_f, \hat{m}_{-f}|\theta_f) > \Pi(\hat{m}_f(\theta_f), \hat{m}_{-f}|\theta_f)$ if $f \in B_2(\hat{m}(\theta))$ and $B_3(\hat{m}(\theta)) = \emptyset$. Then, $\hat{m}$ is not an equilibrium.

Next, suppose there is an equilibrium $\hat{m}$ for which $\hat{m}(\theta) = m$ such that $B_3(m) = \emptyset$ and $f \notin B_2(m)$ for all $f \in F$ has positive probability. Then $B_3(m) = \emptyset$ and $j \in B_2(m)$ with $j \notin F$ has positive probability. Agent j always has a neighbor $f \in F$.

We show that there exist m'_j such that $\Pi(m'_j, \hat{m}_{-j}|\theta_j) > \Pi(m_j, \hat{m}_{-j}|\theta_j)$ and thus $\hat{m}$ is not an equilibrium.

If $B_3(m) = \emptyset$ and $f \notin B_2(m)$, then $\pi_f(m) = 0$ because either the prize is assigned to all agents in $B_2(m)$ or to an agent $k \notin N_j$ if $B_2(m) = \{j\}$. Thus, f must say the truth. As j does not conflict with f it must be true that $m_{jf} = d_f = m_{ff}$ and $m_{jj} = d_j = m_{fj}$. Also j must send a best application, hence $m_{jj} = d_j < d_f = m_{ff}$ since $d_j \neq d_f$.

If $d_f > m_{jj} = d_j > 0$, then j can deviate to $m'_{jj} < d_j$, $m'_{jf} > m_{jj}$ such that $m'_{jf} \neq d_f$ for all f and $m'_{jk} = m_{jk}$ for all $k \neq f$. Then $B_3(m'_j, m_{-j}) = \{j\}$ and $\pi_j(m'_j, m_{-j}) = 1$. If $\pi_j(m) > 0$ and it is not the case that $j \in B_2(m)$ and $B_3(m) = \emptyset$, then either $j \in B_3(m)$ or $B_2(m) = \{k\}$ for $k \notin N_j$ and $B_3(m) = \emptyset$. In these cases deviating to m'_j either leads to $B_3(m'_j, m_{-j}) = \{j\}$ or to $B_2(m'_j, m_{-j}) = \{k\}$ for $k \notin N_j$ and $B_3(m'_j, m_{-j}) = \emptyset$ and thus $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$. Observe that after the deviation j now sends a better application than before and j surely conflicts with f. The certain conflict with f stems from the fact that in each of the two cases, either $m_{ff} \leq m_{jj}$ if $f \in B_3(m)$ or $m_{ff} = d_f$ if $f \notin B_3(m)$ because f only lies if this increases her chances of receiving the prize.

If $m_{jj} = d_j = 0$, then j can deviate to $m'_{jj} = 0$, $m'_{jk} = m_{jk}$ for all $k \neq f$, and $m'_{jf} > b$ such that $m'_{jf} \neq d_f$ for all f. If $j \in B_2(m)$ and $B_3(m) = \emptyset$, then $B_3(m'_j, m_{-j}) = \{j\}$. If $\pi_j(m) > 0$ and not $j \in B_2(m)$ and $B_3(m) = \emptyset$, then either $j \in B_3(m'_j, m_{-j}) \subseteq B_3(m)$ or still $B_2(m'_j, m_{-j}) = \{k\}$ for $k \notin N_j$ and $B_3(m'_j, m_{-j}) = \emptyset$ and thus $\pi_j(m'_j, m_{-j}) \geq \pi_j(m)$. Again

j surely conflicts with f after the deviation because either $m_{ff} = m_{jj}$ if $f \in B_3(m)$ or $m_{ff} = d_f$ if $f \notin B_3(m)$ because f only lies if this increases her chances of receiving the prize.

Hence, if there is a positive probability that $\hat{m}(\theta) = m$ such that $B_3(m) = \emptyset$, $f \notin B_2(m)$ for all f and $j \in B_2(m)$ where j is an agent with partial information, then there always exists m'_j such that $\Pi_j(m'_j, \hat{m}_{-j}|\theta_j) > \Pi_j(\hat{m}_j(\theta_j), \hat{m}_{-j}|\theta_j)$. Then $\hat{m}$ is not equilibrium.

Hence, every equilibrium $\hat{m}$ is such that $B_3(m) \neq \emptyset$ and thus $\pi_i(m) > 0$ if and only if $i \in B_3(m)$ for all $i \in N$ with probability 1.

Next, we show that every equilibrium $\hat{m}$ is such that $B_3(m) = \{g\}$ with probability 1.

Suppose to the contrary that there is an equilibrium $\hat{m}$ in which with positive probability $\hat{m}(\theta) = m$ such that $B_3(m) \neq \{g\}$ where g is $f \in F$. If $B_3(m) \neq \{f\}$, then f can deviate to $m'_{ff} = \max\{0, d_f - b\}$ (the best possible application in the population), and any $m'_{fj} \neq m_{jj}$ if $m'_{ff} > 0$ or $m'_{fj} > d_f + b$ such that $m'_{fj} \neq m_{jj}$ if $m'_{ff} = 0$ for all $j \in N_f$. Then $B_3(m'_f, m_{-f}) = \{f\}$ and hence $\hat{m}$ is not an equilibrium. Thus, every equilibrium $\hat{m}$ is such that with probability 1 $B_3(m) = \{g\}$where $\hat{m}(\theta) = m$ if g is some $f \in F$.

Suppose second that there is an equilibrium $\hat{m}$ in which with positive probability $\hat{m}(\theta) = m$ such that $B_3(m) \neq \{g\}$ where g is local minimum agent $i \notin F$. Observe that every $j \in N_i$ must also be linked to all other $k \in N_i$ and hence knows θ_i because i does not have full information.

If $d_i > b$ and $m_{ii} = 1$, then i can deviate to any m'_i such that $m'_{ii} = d_i - b$ and $m'_{if} = 1$ for all $f \in F$. If i is g and $B_3(m) \neq \{i\}$, then after the deviation i sends the best application and surely conflicts with any $f \in F$. Thus $B_3(m'_i, m_{-i}) = \{i\}$. The conflict with f is certain: If $d_f = 1$ and $m_{ff} = 1$, then the truth is equally good as any lie and f says the truth with $m_{fi} = d_i \neq m'_{ii}$ in any equilibrium. If $d_f = 1$ and $m_{ff} < 1$, then obviously $m'_{if} \neq m_{ff}$. If $d_f < 1$, then the truth is at least as good as any lie with $m_{ff} = 1$ and thus f strictly prefers the truth and sends $m_{ff} < 1 = m'_{if}$ in any equilibrium. In any other case where $i \in B_3(m)$ also $B_3(m'_i, m_{-i}) = \{i\}$ for the same reasons as above.

If $d_i > b$ and $m_{ii} < 1$, then i can deviate to any m'_i such that $m'_{ii} = d_i - b$ and $m'_{if} > m_{ii}$ and $m'_{if} \neq d_f$ for all f. After the deviation, i surely conflicts with any $f \in F$ because in any equilibrium either $m_{ff} \leq m_{ii}$ if $f \in B_3(m)$ or f says the truth with $m_{ff} = d_f$ if $f \notin B_3(m)$. If i is g and $B_3(m) \neq \{i\}$, then $B_3(m'_i, m_{-i}) = \{i\}$. In any other case where $i \in B_3(m)$, then $i \in B_3(m'_i, m_{-i}) \subseteq B_3(m)$.

Hence, $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ in both cases and $\hat{m}$ was not an equilibrium. Thus, every equilibrium $\hat{m}$ is such that with probability 1 $B_3 = \{g\}$ if g is local minimum agent i with partial information and $d_i > b$.

Next assume $d_i \leq b$.

Suppose agent i expects with positive probability that $\hat{m}(\theta) = m$ for which some agent j with $d_j > b$ is in $B_3(m)$. Agent i can only assign a positive probability to this if $m_{ii} > 0$ and $m_{jj} > 0$ for all $j \notin F$ who are in N_i. Then i can deviate to any m'_i such that $m'_{ii} = 0$ and $m'_{if} > \max\{d_i + b, m_{ii}\}$ with $m'_{if} \neq d_f$ for all $f \in F$. After the deviation, i surely conflicts with any $f \in F$ because again either $m_{ff} \leq m_{ii}$ or $m_{ff} = d_f$. If $j \in B_3(m)$ with $d_j > b$, then $B_3(m'_i, m_{-i}) = \{i\}$. In any other case where $i \in B_3(m)$, also $B_3(m'_i, m_{-i}) = \{i\}$. Thus, $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ if $d_i \leq b$ and $\hat{m}$ is not an equilibrium.

Hence, there is no equilibrium $\hat{m}$ in which with positive probability $\hat{m}(\theta) = m$ such that some $j \in B_3(m)$ with $d_j > b$ if j is not g. Thus, every equilibrium $\hat{m}$ is such that with probability 1 j is not in $B_3(m)$ and $\pi_j(m) = 0$ for all such j. Then, every j with $d_j > b$ who is non-minimal says the truth in every equilibrium $\hat{m}$.

Next, we show that every $f \in F$ with $d_f \leq b$ also says the truth in every equilibrium $\hat{m}$. Let $C = \{i, N_i \backslash F\}$. Let $f_1(A)$ be agent $i \in A \subseteq F$ such that $d_i \leq d_k$ for all $k \in A$ and any $A \subseteq F$. Thus $f_1(A)$ is the "best" agent in $A \subseteq F$. Let $D = \{i | i \in F, d_i \leq b\}$.

Suppose $d_f \leq b$ for some $f \in F$.

Suppose $\hat{m}$ is an equilibrium in which with positive probability $\hat{m}(\theta) = m$ such that $m_{kk} > 0$ for all $k \in C$. In this case and if all $k \notin N_i$ have $d_k > b$ and say the truth, then agent $f_1(F)$ chooses m_{f_1} such that $B_3(m) = \{f_1\}$ to maximize π_{f_1}. Then, i can deviate to any m'_i with $m'_{ii} = 0$ and $m'_{if} > \max\{d_i + b, m_{ii}\}$ such that $m'_{if} \neq d_f$ for all f which leads to $B_3(m'_i, m_{-i}) = \{i\}$. In any other case where $i \in B_3(m)$, also $B_3(m'_i, m_{-i}) = \{i\}$. Thus, $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ and $\hat{m}$ is not an equilibrium. Then every equilibrium $\hat{m}$ must be such that $m_{kk} = 0$ for some $k \in C$ with probability 1.

Suppose $\hat{m}$ is an equilibrium in which with positive probability $\hat{m}(\theta) = m$ such that $m_{kk} = 0$ for some $k \in C$, $\max_{l \in C} m_{lf} \leq d_k + b$ for all $f \in A \subseteq D$ and all $k \in C$ who send $m_{kk} = 0$, and $\max_{l \in C} m_{lf} > d_k + b$ for all $f \in D \backslash A$ and some $k \in C$ who sends $m_{kk} = 0$. In words, at least one agent in C sends a zero application, the references agents in C send about agents in $A \subseteq D$ are weakly less than $d_k + b$ for all $k \in C$ who send a zero application, and some reference agents in C send about agent f for all $f \in D \backslash A$ is strictly greater than $d_k + b$ for some $k \in C$ who sends a zero application.

First, every $f \in F \backslash D$ is saying the truth as shown before. Second every $f \in D \backslash A$ is saying the truth because $\bar{r}_f(m_f, m_{-f}) > \min_{k \in B_1(m)} \bar{r}_k(m_f, m_{-f})$ for any m_f. Third, we consider agents in A. Suppose $|A| > 1$ and a subset $L \subseteq A$ with $|L| > 1$ is lying which means that all $l \in L$ must be in $B_3(m)$. Thus $m_{ll} = 0$ and $\bar{r}_l(m) = \min_{k \in B_1(m)} \bar{r}_k(m)$ for all $l \in L$. Then however $f_1(L)$ can deviate to a message m'_{f_1} such that $\bar{r}_l(m'_{f_1}, m_{-f_1}) > \min_{k \in B_1(m'_{f_1}, m_{-f_1})} \bar{r}_k(m'_{f_1}, m_{-f_1})$ for all $l \in L \backslash f_1$ and then $\pi_{f_1}(m'_{f_1}, m_{-f_1}) > \pi_{f_1}(m)$ and $\hat{m}$ is not an equilibrium. Next consider $|L| = 1$ with $L = \{l\}$. If $d_k > b$ for all $k \notin N_i$ and thus all $k \notin N_i$ and all $f \in F \backslash l$ say the truth, then agent l chooses a lie m_l, e.g. $m_{ll} = 0$ and $m_{lj} = \min\{d_j + b, 1\}$ for all $j \neq l$, such that $B_3(m) = \{l\}$. Then however the global minimum agent $i \in C$ can deviate to a message m'_i such that $m'_{ii} = 0$ and $m'_{ij} = \min\{d_j + b, 1\}$ for all $j \in N_i$ and $B_3(m'_i, m_{-i}) = \{i\}$. In any other case where $i \in B_3(m)$, as well $B_3(m'_i, m_{-i}) = \{i\}$ because i surely conflicts with $f \in F$ who sends $m_{ff} \in \{0, d_f\}$ for all $f \in F$. Then $\Pi(m'_i, \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ and

$\hat{m}$ is not an equilibrium. Suppose $|A| = 1$. If $d_k > b$ for all $k \notin N_i$ and thus all $k \notin N_i$ and all $f \in F \backslash A$ say the truth, then agent $f \in A$ chooses lie m_f, e.g. $m_{ff} = 0$ and $m_{fj} = \min\{d_j + b, 1\}$ for all $j \neq f$, such that $B_3(m) = \{f\}$. Then however the global minimum agent $i \in C$ can deviate to a message m_i' such that $m_{ii}' = 0$ and $m_{ij}' = \min\{d_j + b, 1\}$ for all $j \in N_i$ and $B_3(m_i', m_{-i}) = \{i\}$. In any other case where $i \in B_3(m)$, as well $B_3(m_i', m_{-i}) = \{i\}$ because i surely conflicts with $f \in F$ who sends $m_{ff} \in \{0, d_f\}$ for all $f \in F$. Then $\Pi(m_i', \hat{m}_{-i}|\theta_i) > \Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i)$ and $\hat{m}$ is not an equilibrium. Hence, every equilibrium must be such that with probability 1, $|A| = 0$, meaning that $\max_{l \in C} m_{lf} > d_k + b$ for all $f \in D$ and some $k \in C$ who sends $m_{kk} = 0$, and hence all agents $f \in F$ say the truth.

Next, we show that every equilibrium $\hat{m}$ is such that with probability 1 every neighbor $j \in N_l \backslash F$ with $d_j \leq b$ of a local minimum l with partial information says the truth. The reason is that every equilibrium $\hat{m}$ is such that if a local minimum l with partial information has a neighbor $j \in N_l \backslash F$ with $d_j \leq b$, then l sends m_l with $m_{ll} = 0$ and $m_{lj} > d_l + b$ for all $j \in N_l \backslash F$ with $d_j \leq b$ with probability 1.

Any local minimum l who has $d_l \leq b$ and partial information, and all her neighbors $j \in N_l \backslash F$ expect with positive probability that l is g and that $d_k > b$ for all $k \notin N_l$. Keep in mind for the proof that all $f \in F$ and every non-minimal agent k with $d_k > b$ say the truth.

Let $D_l = \{j | j \in \{l, N_l \backslash F\}, d_j \leq b\}$. Only if $|D_l| > 1$, agent l has a neighbor $j \in D_l \backslash l$ who has $d_j \leq b$. Hence, assume $|D_l| > 1$.

Suppose there is an equilibrium $\hat{m}$ such that $\min_{k \in D_l \backslash j} m_{kk} > 0$ for some $j \in D_l \backslash l$. Thus j chooses m_j with $m_{jj} < \min_{k \in D_l \backslash j} m_{kk}$ to maximize her expected probability of being selected. Then, however, l can deviate to $m_{ll}' = 0$ and $m_{lj}' = \min\{d_j + b, 1\}$ for all $j \in N_l$. If l is g and all $k \notin N_l$ have $d_k > b$, then $B_3(m_l', m_{-l}) = \{l\}$. In any other case where $l \in B_3(m)$, as well $B_3(m_l', m_{-l}) = \{l\}$. Then $\Pi(m_l', \hat{m}_{-l}|\theta_l) > \Pi(\hat{m}_l(\theta_l), \hat{m}_{-l}|\theta_l)$ and $\hat{m}$ is not an equilibrium.

Suppose there is an equilibrium $\hat{m}$ in which there is $j \in D_l \backslash l$ such that $\min_{k \in D_l \backslash j} m_{kk} = 0$ and $\max_{h \in D_l \backslash j} m_{hj} \leq d_k + b$ for all $k \in D_l \backslash j$ who send

$m_{kk} = 0$. Thus j chooses some lie with $m_{jj} = 0$ and $m_{jk} \geq \max_{h \in D_l \backslash j} m_{hj}$ for all $k \in D_l \backslash j$ who send $m_{kk} = 0$. Then, however, l can deviate to $m'_{ll} = 0$ and $m'_{lj} = \min\{d_j + b, 1\}$ for all $j \in N_l$. If l is g and all $k \notin N_l$ have $d_k > b$, then $B_3(m'_l, m_{-l}) = \{l\}$. In any other case where $l \in B_3(m)$, as well $B_3(m'_l, m_{-l}) = \{l\}$. Then $\Pi(m'_l, \hat{m}_{-l}|\theta_l) > \Pi(\hat{m}_l(\theta_l), \hat{m}_{-l}|\theta_l)$ and $\hat{m}$ is not an equilibrium.

Thus every equilibrium $\hat{m}$ is such that for all $j \in D_l \backslash l$, $\min_{k \in D_C \backslash j} m_{kk} = 0$ and $\max_{h \in D_l \backslash j} m_{hj} > d_k + b$ for some $k \in D_l \backslash j$ who sends $m_{kk} = 0$. Then, however, $\Pi(m_j, \hat{m}_{-j}|\theta_j) = 0$ for all m_j and all $j \in D_l \backslash l$. Thus all $j \in D_l \backslash l$ say the truth. Hence, if agent l has a neighbor $j \in N_l \backslash F$ with $d_l \leq b$, then l sends m_l with $m_{ll} = 0$ and $m_{lj} > d_l + b$ for all $j \in N_l \backslash F$ with $d_j \leq b$ with probability 1.

Next, we show that every equilibrium $\hat{m}$ is such that with probability 1 every local minimum l who has $d_l \leq b$ and partial information but no neighbor $j \in N_l \backslash F$ with $d_j \leq b$ sends m_l with $m_{ll} = 0$.

Suppose $\hat{m}$ is an equilibrium such that with positive probability a local minimum l who has $d_l \leq b$ and partial information but no neighbor $j \in N_l \backslash F$ with $d_j \leq b$ sends m_l with $m_{ll} > 0$. Then there is a positive probability that some $k \notin N_l$ is a local minimum with $d_k \leq b$ and every $j \notin N_l$ with $j \neq k$ has $d_j > b$. Thus in order to maximize her expected probability of being selected m_k must be such that $m_{kk} < m_{ll}$. Then however, l can deviate to m'_l with $m'_{ll} = 0$ and $m_{lj} = \{d_j + b, 1\}$ for all $j \in N_l$ such that $\Pi(m'_l, \hat{m}_{-l}|\theta_l) > \Pi(\hat{m}_l(\theta_l), \hat{m}_{-l}|\theta_l)$ and $\hat{m}$ is not an equilibrium.

Thus, every equilibrium is such that every local minimum l with partial information and $d_l \leq b$ sends $m_{ll} = 0$ and receives only truthful references from her neighbors. As the global minimum is local minimum with partial information and $d_g \leq b$, $m_{gg} = 0$ and $\bar{r}_g = d_g$.

Next, we show every equilibrium $\hat{m}$ is such that every non-minimal agent $k \notin F$ with $d_k \leq b$ who is not linked to a local minimum with partial information says the truth. Observe that agent j with $d_j = \min_{h \in N_k} d_h$ is in F. Thus agent k expects with probability 1 that either $f \in F$ is g or that a local minimum l with $d_l \leq b$ and partial information is g. If f is g, then $\pi_k(m_k, m_{-k}) = 0$ for all m_k in every equilibrium. If local minimum l with

$d_l \leq b$ and partial information is g, then $\pi_k(m_k, m_{-k}) = 0$ for all m_k in every equilibrium as well because $m_{gg} = 0$ and $\bar{r}_g(m_k, m_{-k}) < \bar{r}_k(m_k, m_{-k})$. Thus $\Pi(m_k, \hat{m}_{-k}|\theta_k) = 0$ for all m_k in every equilibrium and k says the truth.

Thus, every equilibrium is such that if a local minimum $l \notin F$ with $d_l \leq b$ is g, then $m_{kk} = 0$ if and only if k is a local minimum $l \notin F$ with $d_l \leq b$. Every neighbor of local minimum $l \notin F$ with $d_l \leq b$ says the truth. Thus the local minimum $l \notin F$ with $d_l \leq b$ who is g sends $m_{ll} = \min_{k \in N} m_{kk}$ and receives $\bar{r}_l < \bar{r}_k$ for all $k \neq l$ in $B_1(m)$. Then $B_2(m) = \{g\}$. Since $B_3(m) \neq \emptyset$ and $B_3(m) \subset B_2(m)$, $B_3(m) = \{g\}$ in every equilibrium with probability 1. □

Proof of Proposition 5.8.

Let $\hat{m}^l = \hat{m}$. We show for all i and all θ_i that given $\hat{m}_{-i}$, $\hat{m}_i$ is such that $\Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i) \geq \Pi(m_i, \hat{m}_{-i}|\theta_i)$ for all $m_i \in M_i(\theta_i)$ and $\Pi(\hat{m}_i(\theta_i), \hat{m}_{-i}|\theta_i) > \Pi(\theta_i, \hat{m}_{-i}|\theta_i)$ if $\hat{m}_i(\theta_i) \neq \theta_i$.

Let $\hat{m}(\theta) = m$.

First consider $i = 5$. Consider any θ_5 such that 5 is g, thus a local minimum with full information. Then $\pi_5(m) = 1$. Agent 5 only does not say the truth if $d_j - d_5 \leq 2b$ for some $j \in N_5$. Suppose all neighbors say the truth. If 5 deviates to the truth m'_5, then $B_3(m'_5, m_{-5}) = \emptyset$ and $B_2(m'_5, m_{-5}) = \{g\}$. Suppose some neighbor j does not say the truth because $d_j \leq b$ and $d_5 > 0$. If 5 deviates to the truth, then $m_{jj} = 0 < m'_{55} = d_5$. Thus $5 \notin B_3(m'_5, m_{-5})$ but $B_3(m'_5, m_{-5}) \neq \emptyset$. In both cases, $\pi_5(m'_5, m_{-5}) < 1$.

Consider any θ_5 such that 5 is not g. Then $\pi_5(m) = 0$ and $g \in B_3(m)$. Agent g sends $m_{gg} = \max\{0, d_g - b\}$. If $m_{gg} > 0$, then $d_5 - b > m_{gg}$. If $m_{gg} = 0$ and $d_5 \leq b$, then $m_{g5} = d_5 + b > d_g + b$. Thus, $5 \notin B_2(m'_5, m_{-5})$ for any $m'_5 \in M_5(\theta_5)$. Moreover, if $B_2(m'_5, m_{-5}) = \{j\}$ with $j \neq 5$ and $B_3(m'_5, m_{-5}) = \emptyset$, then $\pi_g(m'_5, m_{-5}) = 0$ because $k \notin N_j$ gets the prize. Hence, $\pi_5(m'_5, m_{-5}) = 0$ for every $m'_5 \in M_5(\theta)$.

Consider second $i \neq 5$.

Consider any θ such that i is a local minimum with partial information.

Observe that no neighbor $j \in N_i$ is in $B_2(m)$ because either $m_{ii} < m_{jj}$, or $m_{ii} = m_{jj} = 0$ but $m_{ij} = d_j + b > d_i + b$.

Suppose i is g.

If $d_j > b$ for $j \notin N_i$, then $B_3(m) = \{i\}$ because $m_{ii} = \min_{k \in N} m_{kk}$, $\overline{r}_i(m) = \min_{k \in B_1(m)} \overline{r}_k(m)$ and i conflicts with neighbor $j \in N_i$ if $m_{ij} - m_{ii} \leq 2b$. Thus $\pi_i(m) = 1$.

If $d_j \leq b$ for $j \notin N_i$ and j says the truth, then $B_3(m) = \{i\}$ and hence $\pi_i(m) = 1$.

If $d_j \leq b$ for $j \notin N_i$ and j does not say the truth, then $m_{ii} = m_{jj} = 0$ for $j \notin N_i$. If all $j \in N_i$ send truthful references, then $B_3(m) = \{i\}$ and hence $\pi_i(m) = 1$. If some $j \in N_i$ sends references equal to b about i and $j \notin N_i$, then $B_3(m) = \{i, j\}$ for $j \notin N_i$ and $\pi_i(m) = \frac{1}{2}$. In the latter case, there does not exist $m_i' \neq m_i$ such that $j \notin B_3(m_i', m_{-i})$ for $j \notin N_i$ because i can neither influence the application of $j \notin N_i$ nor the references she and $j \notin N_i$ receive. Thus there is no $m_i' \neq m_i$ such that $\pi_i(m_i', m_{-i}) > \pi_i(m)$.

Suppose i is not g. This means $j \notin N_i$ is g because the neighbors of i are non-minimal.

If $d_i > b$, or if $d_i \leq b$ and all $j \in N_i$ send $m_{ji} = d_i$ and $m_{jg} = d_g$, then $B_3(m) = \{g\}$ and $\pi_i(m) = 0$ and there does not exist $m_i' \neq m_i$ such that $B_3(m_i', m_{-i}) \neq \{g\}$. If $d_i \leq b$ and some $j \in N_i$ sends $m_{ji} = m_{jg} = b$, then $B_3(m) = \{g, i\}$ and $\pi_i(m) = \frac{1}{2}$ and there does not exist $m_i' \neq m_i$ such that $g \notin B_3(m_i', m_{-i})$. Hence, in each case, there exists no $m_i' \neq m_i$ such that $\pi_i(m_i', m_{-i}) > \pi_i(m)$.

Thus, $\Pi(m_i, \hat{m}_{-i}|\theta_i) \geq \Pi(m_i', \hat{m}_{-i}|\theta_i)$ for all $m_i' \in M_i(\theta_i)$.

Next, we show that if $m_i \neq \theta_i$, then $\Pi(m_i, \hat{m}_{-i}|\theta_i) > \Pi(\theta_i, \hat{m}_{-i}|\theta_i)$.

Agent i with $d_i > 0$ expects with strictly positive probability that i is g and $j \notin N_i$ a local minimum with $d_j \in (d_i, \min\{d_i + b, 1\})$. In this case, $m_{ii} = \max\{0, d_i - b\}$ and $i \in B_3(m)$ and $\pi_i(m) > 0$. If i deviates to the truth with $m_{ii}' = d_i$, then $i \notin B_1(m_i', m_{-i})$ and $B_3(m_i', m_{-i}) = \{j\}$ with $j \notin N_i$ and thus $\pi_i(m_i', m_{-i}) = 0$.

Agent i with $d_i = 0$ only does not say the truth if $d_j - d_i \leq 2b$ for some $j \in N_i$. As $d_i = 0$, all $j \in N_i$ say the truth, and hence $B_3(m) = \{i\}$

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 $\pi_i(m_i', m_{-i}) = 0$.

Consider any θ such that i is non-minimal and has $d_i > b$ or $d_j = 0$ for some $j \in N_i$ or $d_j > b$ for all $j \in N_i \backslash 5$. Then $m_i = \theta_i$, $i \notin B_3(m) \neq \emptyset$ and thus $\pi_i(m) = 0$.

There is no $m_i' \neq m_i$ such that $i \in B_2(m_i', m_{-i})$. If $d_i > b$, then $m_{gg} < d_i - b \leq m_{ii}'$ for any m_i'. If $d_j = 0$ for $j \in N_i$, then j is g and $m_{gg} = 0$ and $m_{gi} > d_g + b \geq m_{ig}'$ for any m_i'. If $d_i \leq b$ and $d_j > b$ for all $j \in N_i \backslash 5$, then either 5 or $j \notin N_i$ is g and $B_2(m_i', m_{-i}) = \{g\}$ for any $m_i' \neq m_i$.

There is no $m_i' \neq m_i$ such that $B_3(m_i', m_{-i}) = \emptyset$ and $B_2(m_i', m_{-i}) = \{j\}$ for $j \notin N_i$. If $j \notin N_i$ is g, then $B_3(m_i', m_{-i}) = \{j\}$ for all $m_i' \neq m_i$. If $j \notin N_i$ is not g, then $j \notin B_2(m_i', m_{-i})$ for any $m_i' \neq m_i$.

There is no $m_i' \neq m_i$ such that $B_3(m_i', m_{-i}) = \emptyset$ and $B_2(m_i', m_{-i}) = \{5\}$. If 5 is g, then $B_3(m_i', m_{-i}) = \{5\}$ for all $m_i' \neq m_i$. If 5 is not g, then $5 \notin B_1(m_i', m_{-i})$ for any $m_i' \neq m_i$.

Thus, $\Pi(\theta_i, \hat{m}_{-i}|\theta_i) \geq \Pi(m_i', \hat{m}_{-i}|\theta_i)$ for all $m_i' \in M_i(\theta_i)$.

Consider any θ such that i is non-minimal and has $d_i \leq b$ and $d_j > 0$ for all $j \in N_i$ and $d_j \leq b$ for some $j \in N_i \backslash 5$.

Suppose $g \notin N_i$. If $d_g > 0$, then some $j \in N_i \backslash 5$ sends $m_{ji} = m_{jg} = b$ because $d_g < d_j \leq b$ 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 $\pi_i(m) \geq \pi_i(m_i', m_{-i})$. If $d_g = 0$, 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 all $j \in N_i$ say the truth.

Suppose $g \in N_i$ and hence $d_g > 0$. Then $i \notin B_3(m) \neq \emptyset$ and $\pi_i(m) = 0$. There is no $m_i' \neq m_i$ such that $i \in B_2(m_i', m_{-i})$ because $m_{gg} = 0$ and $m_{gi} = d_i + b$. 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 $m_{gj} > d_g + b$. Finally, there is no $m_i' \neq m_i$ such that $B_2(m_i', m_{-i}) = \{5\}$ and $B_3(m_i', m_{-i}) = \emptyset$ because either 5 is g and conflicts with $j \in N_5 \backslash i$ or 5 is not g and $m_{gg} = 0$ and $m_{g5} > d_5 + b$.

Thus, $\Pi(m_i, \hat{m}_{-i}|\theta_i) \geq \Pi(m_i', \hat{m}_{-i}|\theta_i)$ for all $m_i' \in M_i(\theta_i)$.

Next, we show that $\Pi(m_i, \hat{m}_{-i}|\theta_i) > \Pi(\theta_i, \hat{m}_{-i}|\theta_i)$. 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 $m'_{ii} > m_{gg}$.

References

3D Group (2016), *Current Practices in 360 Degree Feedback*, 5 edn, 3D Group, Emeryville, CA.

3D Group (2018), '3d group — 360 feedback & consulting services'.
URL: *https://3dgroup.net/*

Aziz, H., Lev, O., Mattei, N., Rosenschein, J. S. and Walsh, T. (2016), 'Strategyproof peer selection using randomization, partitioning, and apportionment', *arXiv preprint arXiv:1604.03632* .

Baumann, L. (2017), 'Identifying the best agent in a network', *SSRN Working Paper* .

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

Bloch, F. and Olckers, M. (2018), 'Friend-based targeting', *Working Paper* .

Cornerstone OnDemand (2018), *Cornerstone OnDemand Announces Fourth Quarter and Fiscal Year 2017 Financial Results.*
URL: *http://investors.cornerstoneondemand.com/investors/news-and-events/news/news-details/2018/Cornerstone-OnDemand-Announces-Fourth-Quarter-and-Fiscal-Year-2017-Financial-Results/default.aspx*

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

Dutta, B. and Sen, A. (2012), 'Nash implementation with partially honest individuals', *Games and Economic Behavior* **74**(1), 154–169.

Dziubiński, M., Sankowski, P. and Zhang, Q. (2016), Network elicitation in adversarial environment, *in* 'International Conference on Decision and Game Theory for Security', Springer, pp. 397–414.

Expert Training Systems (2018), 'Ets – hr consultancy, 360 degree feedback and employee surveys'.
URL: *https://www.etsplc.com/*

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

Green, J. R. and Laffont, J.-J. (1986), 'Partially verifiable information and mechanism design', *The Review of Economic Studies* **53**(3), 447–456.

Jackson, M. O. (1991), 'Bayesian implementation', *Econometrica: Journal of the Econometric Society* pp. 461–477.

Jackson, M. O., Palfrey, T. R. and Srivastava, S. (1994), 'Undominated Nash implementation in bounded mechanisms', *Games and Economic Behavior* **6**(3), 474–501.

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

Koessler, F. and Perez-Richet, E. (2017), 'Evidence reading mechanisms', *Working Paper* .

Korpela, V. (2014), 'Bayesian implementation with partially honest individuals', *Social Choice and Welfare* **43**(3), 647–658.

Kurokawa, D., Lev, O., Morgenstern, J. and Procaccia, A. D. (2015), Impartial peer review, *in* 'IJCAI', pp. 582–588.

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

Maskin, E. (1999), 'Nash equilibrium and welfare optimality', *The Review of Economic Studies* **66**(1), 23–38.

Merrifield, M. R. and Saari, D. G. (2009), 'Telescope time without tears: a distributed approach to peer review', *Astronomy & Geophysics* **50**(4), 4–16.

Nierstrasz, O. (2000), 'Identify the champion', *Pattern Languages of Program Design* **4**, 539–556.

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

Saba Software (2018), 'About — saba software'.
URL: *https://www.saba.com/uk/about*

Strausz, R. (2016), 'Mechanism design with (partially) verifiable information and the revelation principle', *Working Paper* .