Friend-based Ranking

Francis Bloch* Matthew Olckers†

October 17, 2018

Abstract

We analyze the design of a mechanism to extract ordinal information disseminated in a social network. We show that friend-based ranking—the report by agents on the characteristics of their neighbors—is a necessary condition for ex post incentive-compatible and efficient mechanism design. We characterize the windmill network as the sparsest social network for which the planner can construct a complete ranking. When complete rankings cannot be achieved, ex post incentive-compatible and efficient mechanisms arise when social networks are bipartite or composed of triangles. We illustrate these findings using real social networks in India and Indonesia.

KEYWORDS: social networks, mechanism design, peer ranking, targeting

JEL CLASSIFICATION NUMBERS: D85, D82, O12, D71

We thank Abhijit Banerjee, Leonie Baumann, Bhaskar Dutta, Matt Elliott, Ben Golub, Sanjeev Goyal, Matt Jackson, Matt Leister, Neil Lloyd, Brian Rogers, Jonathan Weinstein, Leeat Yaariv and Yves Zénou for helpful discussions on the project. We are also grateful to seminar participants at several institutions for their comments. This work was partly supported by the Agence Nationale de la Recherche under grant ANR13BSHS10010.

*Paris School of Economics and Université Paris 1 Panthéon-Sorbonne, francis.bloch@univ-paris1.fr

†Paris School of Economics and Université Paris 1 Panthéon-Sorbonne , matthew.olckers@psemail.eu

1 Introduction

In many social networks, individuals gradually acquire information about their neighbors through repeated interactions. Pupils in a class learn about the ability of other pupils with whom they write joint projects, workers in a firm learn about the productivity of the coworkers in their teams, and members of a community in a developing country learn about the needs of their close friends. This information, which is disseminated in the social network, may be of great use to an external planner who wants to extract information about members of the community. A teacher wants to learn about the ability of her pupils; an employer, the productivity of her workers; a funding agency, the needs of villagers in a developing country.

In the classical literature on mechanism design, the principal designs a mechanism which asks individuals to report on their own types. However, a large number of mechanisms is used in practice which ask individuals to report not on their own type but on the type of others. Pupils are asked to assess the performance of other pupils, workers are asked to measure the productivity of their coworkers, or villagers are asked to rank other individuals in the community. The objective of our paper is to analyze these mechanisms, that ask individuals to report about their neighbors in the social network—mechanisms that we term "friend-based ranking mechanisms." In particular, we want to understand how the architecture of the fixed social network affects the planner's ability to construct a mechanism having desirable properties.

While we treat the social network as exogenous, we note that in some situations the planner can design the network endogenously. Consider, for example, peer selection problems, e.g., when editors of scientific journals ask scholars to review the work of their peers, or scientific funding agencies ask applicants to review the projects of other applicants. By assigning reviewers to projects, the planner designs a network of observation which plays exactly the same role as the exogenous social networks described above.

We study a setting with three characteristics. First, we assume that information is local. An individual may make comparisons only among his direct neighbors. Second, information is ordinal. Individuals lack the ability to quantify characteristics and can only assess whether one individual has a higher characteristic than another. Third, we assume that the planner has only one instrument at her disposal: she constructs a (complete) ranking of the members of the community. Hence, the number of outcomes that the planner can select from is very restricted. The planner cannot use transfers, and cannot punish individuals by excluding them from the ranking. In particular, she cannot impose punishments for inconsistent reports, as in the classical literature on implementation with correlated types (Crémer and McLean, 1985).

We require that the planner's mechanism satisfy two properties. First, individuals must have an incentive to report information truthfully. In the ordinal setting we consider, the natural choices for implementation concepts are dominant-strategy and ex post implementation. However, we notice that the limited number of outcomes in our setting means that dominant-strategy implementation is too strong, leading to impossibility results. We adopt instead ex post incentive compatibility as the desirable incentive property of the mechanism.

Second, we require the mechanism to be ex post efficient from the point of view of the planner, whose objective is to recover the true ranking of individuals in the community. More precisely, the ranking chosen by the planner must match the ranking that society would construct by aggregating all local information. If society can construct a complete ranking of individuals for any realization of types (a situation we label "completely informative"), the ranking chosen by the planner must match the true ranking. Otherwise, we may face different situations according to the realization of types. For some type realizations, even if the information aggregated by society is not complete, transitivity ensures that all individuals can be completely ranked. For other type realizations, society will be able to construct only a partial order on the individuals. In the latter case, the complete ranking chosen by the planner will be a completion of the partial order that the community is able to construct, and this completion will involve an arbitrary ranking across individuals who cannot be compared.

We first analyze mechanism design in completely informative societies. We show that a society is *completely informative* if every pair of individuals can be compared, either through "self-reports" (the two individuals involved in the pair report on each other) or through "friend-based reports" (a third individual observes both individuals in the pair). Our main theorem shows that self-reports can be used only if they are backed by the report of a third individual. A mechanism satisfying ex post incentive compatibility and efficiency exists if and only if every pair of individuals has a common neighbor. We then characterize the sparsest network which satisfies this property. When the number of individuals is odd, this is the "friendship network" of Erdős, Rényi, and Sós (1966), the only network in which every pair of individuals has only one common friend. This network, also known as the windmill, has one individual as a hub who connects all other individuals who form pairs. When the number of individuals is even, this is a variant of the windmill in which one of the "sails" contains three individuals instead of two. In this network, one individual—the hub— is responsible for a large number of comparisons, so we also investigate other networks where every pair of agents has a common neighbor when the number of comparisons performed by every individual is capped.

We then turn our attention to societies which are not completely informative; to study these, we add another requirement to the mechanism: to guarantee that whenever two individuals are incomparable, the mechanism ranks them in the same way independent of (irrelevant) information from other comparisons. Under this independence requirement, we show that any comparison based solely on self-reports must be discarded by the planner, as both individuals have an incentive to misreport. Hence, the planner can rely only on friend-based comparisons, and we construct a "comparison network" by linking two individuals if and only if they have a common neighbor. We find that there exist two network architectures for which the planner can construct a mechanism satisfying independence, ex post incentive compatibility, and efficiency. In the first architecture, the social network is bipartite (which is equivalent to the comparison network being disconnected). We use the bipartite structure to partition the set of individuals, so that individuals in one group rank individuals in the other group, and individuals are ranked across groups in an arbitrary way. In the second architecture, all links form triangles, and we can construct a mechanism exploiting the fact that any unsupported report is surrounded by supported links. However, we also note that there exist network architectures for which mechanisms satisfying all three properties named above cannot be constructed. The simplest example is a network of four individuals with one triangle and one additional link.

We then highlight three aspects of our findings using social network data from villages in Karnataka, India, and neighborhoods in Indonesia. First, information, as measured by the share of unique comparisons the planner receives, depends on network structure. Two social networks of similar density may provide very different levels of information. Second, we decompose comparisons into those within a triangle, those across triangles, and a remainder. Low-density networks have a large share of across-triangle comparisons. Third, we simulate the process of capping the number of comparisons each individual provides. If the cap is small relative to the community size, the capped network is close to the upper bound in information (as measured by the number of unique comparisons).

Finally, we consider different variants and robustness checks of our model. We show that dominant-strategy implementation is too strong, leading to an impossibility result in triangles. We analyze the robustness of our mechanism to joint deviations by groups. We study whether coarser rankings are easier or harder to implement than complete rankings. We study the impact of homophily. If individuals of similar characteristics are more likely to form friendships, the planner is more likely to extract the necessary and sufficient friend-based comparisons to find the complete ranking. This relationship is reversed when the probability of across-group links is close to zero.

In practical terms, our analysis points to two important facts. First, it shows that it may be useful to partition the set of individuals into different groups and ask individuals in one group to rank those in another. For example, one may want to let men rank women and women rank men in a community. This procedure will result in a truthful and efficient ranking, but the price to be paid is that interrankings among individuals in the two groups will be arbitrary. Second, our analysis highlights the importance of triangles. Truthful and efficient comparisons will be possible if all links form triangles, suggesting that friend-based ranking should be used only in societies with high clustering. As high clustering is also associated with high density and low average distances, we conclude that friend-based ranking should be used only in communities with dense social networks with low diameters. Finally, our analysis can be used to help design review systems in peer-selection problems. It suggests that using asymmetric networks of observation, with central reviewers observing a large number of projects, may be a way to construct efficient and incentive-compatible peer-review systems.

1.1 Literature Review

We first discuss the relationship of our paper with the literature on community-based ranking in development economics. This literature documents experiments in which members of a community are gathered to collectively agree on a ranking in order to identify the poorest or the most able individuals. Rai (2002) discusses the individual incentives to lie in poverty targeting. Alatas, Banerjee, Hanna, Olken, and Tobias (2012) and Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2016) report on an experiment in Indonesia in which community members were asked to collectively identify recipients of benefits of social programs. They compare the accuracy of community-based targeting with traditional proxy-means testing, and argue that community-based targeting results in consensus, and brings higher satisfaction to all members of the community. Hussam, Rigol, and Roth (2017) report on a recent field experiment in Maharashtra, India where entrepreneurs were asked to rank their peers according to the profitability of their businesses. They prove that this is a more accurate method of ranking than using observable information about the entrepreneurs. We see friend-based ranking as a complement to targeting methods that are currently popular such as proxy-means tests. Although widely used, the proxy means test has been shown to perform only slightly better than universal transfers at reducing poverty (Brown, Ravallion, and Van de Walle, 2016) and to lack adjustment to transitory shocks (Coady, Grosh, and Hoddinott, 2004).

The theoretical analysis of the paper is closely related to the limited literature in computer science and social choice theory studying peer selection. Alon, Fischer, Procaccia, and Tennenholtz (2011) analyze the design of mechanisms to select a group of k individuals among their peers. Alon, Fischer, Procaccia, and Tennenholtz (2011) prove a strong negative result: no deterministic efficient strategy-proof mechanism exists. Approximately efficient, stochastic, impartial mechanisms can be constructed, which are based on the random partition of individuals into clusters of fixed size such that individuals inside a cluster rank individuals outside the cluster. Holzman and Moulin (2013) analyze impartial voting rules when individuals nominate a single individual for office. They identify a class of desirable voting rules as two-step mechanisms, by which voters are first partitioned into districts which elect local winners, who then compete against one another to select the global winner. However, Holzman and Moulin (2013) also highlight a number of impossibility results, showing that there is no impartial voting procedure which treats voters symmetrically, nor any impartial voting procedure which guarantees (i) that an individual whom nobody considers best will be elected and (ii) that an individual whom everybody considers best will always be elected. Kurokawa, Lev, Morgenstern, and Procaccia (2015) and Aziz, Lev, Mattei, Rosenschein, and Walsh (2016) improve on the partition algorithm proposed in Alon, Fischer, Procaccia, and Tennenholtz (2011). They consider a more general setting, inspired by the new peer-review system instituted by the National Science Foundation to fund the Sensors and Sensing System program. Kurokawa, Lev, Morgenstern, and Procaccia (2015) propose the “credible subset mechanism,” a process which first identifies candidates who are likely to win, and assigns ratings only to these candidates. Aziz, Lev, Mattei, Rosenschein, and Walsh (2016) propose a mechanism combining the insights of the partition mechanism of Alon, Fischer, Procaccia, and Tennenholtz (2011) with the impartial “divide the dollar” mechanism of De Clippel, Moulin, and Tideman (2008).

Our model departs from all these models of peer selection in a number of ways. First, we consider *ordinal* rather than cardinal information as inputs to the mechanism. In our model, individuals do not assign grades to other individuals, but can only make bilateral comparisons. Second, we consider as output a *complete ranking* of individuals rather than a coarse ranking into two sets of acceptable and non acceptable candidates. (However, in Section 6, we also consider coarser rankings as a possible extension of our model.) Third, because dominant-strategy mechanisms do not exist, we weaken the incentive requirement to ex post implementation, thereby obtaining positive results which differ from the results obtained in the peer-selection literature. Fourth, and most importantly, we do not assume a specific assignment of proposals to reviewers, but consider an arbitrary network of observations captured by a social network. Our main objective is then to characterize those social networks (or structures of observability) for which mechanisms satisfying desirable properties can be constructed.

The paper which is probably the more closely connected to ours is a recent paper by Baumann (2017) which analyzes network structures for which it is possible to identify the individual with the highest characteristic. Baumann (2017) constructs a specific multitier mechanism identifying the top individual from the reports of his neighbors. The mechanism admits multiple equilibria, but there are some social network architectures (e.g., the star) for which all equilibria result in the identification of the top individual. Our paper differs from Baumann's, however, in many dimensions. First, we consider an ordinal rather than a cardinal setting, giving rise to the possibility of incompleteness of the social ranking. Second, we assume that the objective of the planner is to rank all individuals rather than identify the top individual. Third, we do not assume an exogenous bound on the way in which individuals can misreport, in contrast to Baumann (2017), in which this exogenous bound plays a crucial role in the construction of equilibria.

2 Model

2.1 Individuals and communities

We consider a community N of n individuals indexed by $i = 1, 2, ..., n$. Each individual i has a characteristic $\theta_i \in \mathbb{R}$. Examples of θ_i include wealth, aptitude for a job, or quality of a project. Characteristics are privately known and are drawn from a nonatomic continuous distribution F.

Members of the community are linked by a connected, undirected graph g. The social network g is common knowledge among the individuals and the planner. The characteristic of individual i, θ_i, can be observed by individual i and by all his direct neighbors in the social network g. We suppose that individuals cannot provide an accurate value for the characteristic θ_i. Either the characteristic cannot be measured precisely, or individuals do not have the ability or the language to quantify θ_i precisely. Instead, we assume that individuals possess *ordinal* information and are able to compare the characteristics of two individuals. For any individual i and any pair of individuals (j, k) that individual i can observe, we let $t^i_{jk} = 1$ if individual i observes that $\theta_j > \theta_k$, and $t^i_{jk} = -1$ if individual i observes that $\theta_j < \theta_k$. The ordinal comparison is assumed to be perfect: individual i always perfectly observes whether individual j's characteristic is higher than that of individual k. Given that the characteristics are drawn from a nonatomic continuous distribution, we ignore situations in which the two characteristics are equal.

Individual i's information (and type) can thus be summarized by a matrix $T^i = [t^i_{jk}]$, where $t^i_{jk} \in \{-1, 0, 1\}$ and $t^i_{jk} \neq 0$ if and only if i observes the comparison between j and k, namely either $i = j$ or $i = k$ or $g_{ij}g_{ik} = 1$. When $i = j$ or $i = k$, we call the comparison t^i_{jk} a *self-comparison*. When $g_{ij}g_{ik} = 1$, we call the comparison t^i_{jk} a *friend-based comparison*.

The vector $\mathbf{T} = (T^1, .., T^n)$ describes the information possessed by the community on the ranking of the characteristics of all the individuals. Obviously, because individual observations are perfectly correlated, individual types T^i and T^j will be correlated if there exists a pair of individuals (k, l) such that $t^i_{kl} \neq 0$ and $t^j_{kl} \neq 0$. Hence, if the planner could construct a punishment for contradictory reports, as in

Crémer and McLean (1985), she would be able to induce the individuals to report their true type. However, we rule out arbitrary punishments.

The information contained in the vector $\mathbf{T} = (T^1, .., T^n)$ results in a partial ranking of the characteristics of the individuals, which we denote by $\succ$. We let $i \succ_{\mathbf{T}} j$ if the information contained in $\mathbf{T}$ allows us to conclude that $\theta_i > \theta_j$.

For a fixed social network g, the information contained in the vector $\mathbf{T} = (T^1, .., T^n)$ may not be the same for different realizations of $(\theta^1, .., \theta^n)$. This is due to the fact that (i) new comparisons can be obtained by transitivity but (ii) the transitive closure of an order relation depends on the initial order relation. To illustrate this point, consider four individuals $i = 1, 2, 3, 4$ organized in a line as in Figure 1

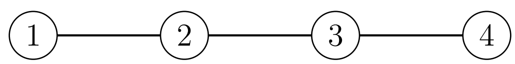


Figure 1: A line of four individuals

If $\theta_1 < \theta_2 < \theta_3 < \theta_4$, then given that $t^1_{12} = t^2_{12} = -1, t^2_{23} = t^3_{23} = -1, t^3_{34} = t^4_{34} = -1, t^2_{13} = -1$, and $t^3_{24} = -1$, the comparisons result in a complete ranking $1 \prec 2 \prec 3 \prec 4$. However, for other possible realizations of $(\theta_1, \theta_2, \theta_3, \theta_4)$, the ranking generated by the types $\mathbf{T}$ may be incomplete. For example, if $\theta_1 < \theta_4 < \theta_2 < \theta_3$, we obtain $1 \prec 2 \prec 3$ and $4 \prec 2 \prec 3$, but 1 and 4 cannot be compared.

A social network g is called *completely informative* if, for any realization of the characteristics $(\theta_1, .., \theta_n)$, the information contained in $\mathbf{T}$ results in a complete ranking of the members of the community. The following lemma characterizes completely informative social networks.

Lemma 1. *A social network g is completely informative if and only if, for any pair of individuals (i, j) either $g_{ij} = 1$ or there exists an individual k such that $g_{ik}g_{jk} = 1$.*

A social network is completely informative if and only if *every pair of individuals* can be compared either by self-comparisons or by friend-based comparisons.

2.2 Planner and mechanism design

The objective of the planner is to construct a ranking of individuals according to the value of the characteristic θ_i. For example, a charity wishes to rank potential beneficiaries by need, an employer wants to rank workers according to their ability, a bank wants to rank projects according to their profitability. We let ρ denote the complete order chosen by the planner. The set of all complete orders is denoted by $\mathcal{P}$. The rank of individual i is denoted by ρ_i.

The planner wishes to construct a ranking as close as possible to the true ranking of the values of the characteristic θ_i. We do not specify the preferences of the planner. In the ordinal setting that we consider, different measures of distances between rankings can be constructed. Instead of describing explicitly the loss function associated with differences in rankings, we focus attention on efficient mechanisms. Efficiency requires that the ranking ρ coincides with the ranking generated by $\mathbf{T}$ for any pair of individuals (i, j) who can be compared under $\mathbf{T}$.

Individuals care only about their rank ρ_i and have strict preferences over ρ_i. By convention, individuals prefer higher values of the ranking. Hence, ρ_i is preferred to ρ'_i if and only if $\rho_i > \rho'_i$. In particular, we assume that there are no externalities in the community, and thus, individuals do not derive any reward from high rankings of friends or low rankings of foes.

A *direct mechanism* associates to any vector of reported matrices $\mathbf{T} \in \mathcal{T}^n$ a complete ranking $\rho \in \mathcal{P}$. We impose the following two conditions on the mechanism:
Ex post incentive compatibility. For any individual i, for any vector of types $\mathbf{T} = (T^i, T^{-1})$, any type $T^{'i}$, the following holds

$$\rho_i(\mathbf{T}) \geq \rho_i(T^{'i}, T^{-i}).$$

Ex post efficiency. For any vector of types $\mathbf{T}$, and for any pair of individuals i and j, the following holds

$$\text{if } i \succ_{\mathbf{T}} j, \text{ then } \rho_i(\mathbf{T}) > \rho_j(\mathbf{T}).$$

We focus on ex post implementation for two reasons. First, because we consider an ordinal setting, we select a robust implementation concept which does not depend on the distribution of types. Second, as we show in section 6, the alternative robust implementation concept—dominant-strategy implementation— is too strong for our setting.

Ex post efficiency requires that the planner's ranking coincide with the true ranking of characteristics in a very weak sense. Whenever two individuals i and j can be ranked using the information contained in $\mathbf{T}$, the ranking ρ_i must be consistent with the ranking between i and j. As the order relation induced by $\mathbf{T}$, $\succ_{\mathbf{T}}$, may be very incomplete, the requirement may be very weak. The ranking ρ must be a completion of the ranking $\succ_{\mathbf{T}}$. If $\succ_{\mathbf{T}}$ is a very small subset of N^2, the ranking ρ may end up being very different from the true ranking of the values of the characteristic θ_i. However as the true ranking of characteristics cannot be constructed using the local information from the social network, the difference between ρ and the true ranking should not be a matter of concern, since the planner chooses an efficient mechanism given the information available to the community.

3 Completely informative rankings

3.1 Importance of common friends

We first analyze conditions under which an ex post incentive-compatible and efficient mechanism can be constructed when the information available in the community results in a complete ranking. By Lemma 1, all pairs of individuals must either be directly connected, or observed by a third individual. The next theorem shows that for an ex post incentive-compatible and efficient mechanism to exist, all pairs of individuals must be observed by a third individual.

Theorem 1. *Suppose that the social network g is completely informative. An ex post incentive-compatible and efficient mechanism exists if and only if, for all pair of individuals (i, j), there exists a third individual k who observes both i and j, i.e., $g_{ik}g_{jk} = 1$.*

Theorem 1 shows that an ex post incentive-compatible and efficient mechanism exists in completely informative communities if and only if every pair of individuals (i, j) has a common friend k. Self-comparisons cannot be used. Every comparison requires the presence of a third party. If the two individuals i and j are connected, the link ij must be 'supported' by a third individual, following the terminology of Jackson, Rodriguez-Barraquer, and Tan (2012).

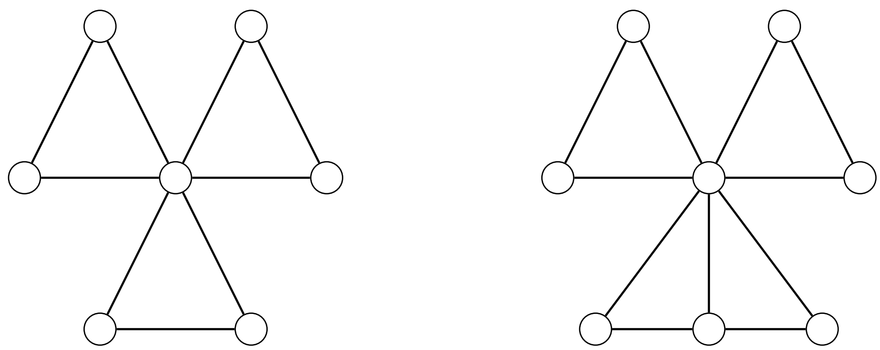
The intuition underlying Theorem 1 is easy to grasp. If the comparison between θ_i and θ_j can be reported only by i and j, in an ex post efficient mechanism, one of them has an incentive to lie. Consider a ranking which places i and j as the two individuals with the lowest characteristics in the community. If both announce that θ_i is smaller than θ_j, then $\rho_i = 1, \rho_j = 2$. Similarly, if both announce that θ_j is smaller than θ_i, then $\rho_j = 1, \rho_i = 2$. But by incentive compatibility, neither of the individuals can improve his rank by changing his report on t_{ij}. Hence i must still be ranked at position 1 when she announces $\theta_i > \theta_j$ and j announces $\theta_i < \theta_j$, and similarly individual j must still be ranked at position 1 when she announces $\theta_j > \theta_i$ and individual i announces $\theta_i > \theta_j$. As two individuals cannot occupy the same position in the ranking, this contradiction shows that there is no ex post incentive-compatible and efficient mechanism relying on self-reports. Notice that this impossibility result stems from the fact that the planner has a very small number of outcomes at her disposal. If she could impose any arbitrary punishment (for example by excluding all individuals who provide inconsistent reports), she could implement an ex post efficient mechanism in dominant strategies, as in Crémer and McLean (1985), for any network architecture.

The construction of an ex post incentive-compatible and efficient mechanism when all links are supported is very intuitive. First consider a comparison between i and j which is observed by at least three individuals. The mechanism disregards the report of any individual who deviates from the reports of all other individuals. Hence no individual can unilaterally change the outcome of the mechanism when all other individuals report the truth. Next suppose that the comparison between i and j is dictated by a third party, a common friend k of i and j. A change in reports could not improve the rank of k given that all other individuals tell the truth and that the social network is completely informative. If the change in report creates an inconsistency in the ranking, the planner can detect if a single individual has cheated and punish him by ranking him at the worst position in the ranking. If the change in report does not create a violation in transitivity, because the social network is completely informative, the rank of individual k is fully determined by the reports of other individuals in the community. The rank of individual k is fixed and no change in report can improve the position of individual k in the ranking. This "friend-based" ranking mechanism is ex post incentive-compatible and efficient.

Theorem 1 characterizes communities for which friend-based ranking mechanisms can be constructed. Clearly the complete network satisfies the conditions. However, the condition is also satisfied by many other social networks, which are less dense than the complete network. Our next result characterizes the *sparsest* networks for which the condition of Theorem 1 holds. This characterization is based on the "friendship theorem" of Erdős, Rényi, and Sós (1966).

Theorem 2. *(The "friendship theorem") If G is a graph of order n in which any two* *vertices i and j have one neighbor in common, then $n = 2m + 1$ and G contains m triangles which are connected at a common vertex.*

The "friendship theorem", initially stated and proved in Erdős, Rényi, and Sós (1966), asserts that in any community where every pair of individuals has exactly one friend in common, one individual is friends with everyone and is the common friend of all other individuals.[1] The "friendship graph" is illustrated in Figure 2 for $n = 7$. For obvious reasons, it is also called the "windmill graph". The friendship graph has exactly $3m$ edges. Our next theorem shows that this is actually the smallest number of edges for which a completely informative mechanism can be constructed when n is odd. When n is even, the graph which minimizes the number of edges is a variation of the friendship graph, where one of the sails of the windmill contains three vertices, as illustrated in Figure 2 for $n = 8$.



(i) $n = 7$. For odd number of nodes, the windmill is also called a friendship graph.

(ii) $n = 8$. For even number of nodes the windmill is modified and one sail has three nodes.

Figure 2: Windmill graphs

Theorem 3. *Suppose that $n \geq 3$. Let g be a social network for which friend-based ranking generates a complete ranking. Then g must contain at least $\frac{3n}{2} - 1$ links if n is even and $\frac{3(n-1)}{2}$ links is n is odd. If n is odd, the unique sparsest network architecture is the friendship network. If n is even, the unique sparsest network architecture is a modified windmill network where one of the sails contains three nodes i, j, k such that i, j and k are connected to the hub, i is connected to j and j is connected to k.*

[1]Different proofs of the friendship theorem have been proposed, often using complex combinatorial arguments (Wilf, 1971; Longyear and Parsons, 1972; Huneke, 2002).

Theorem 3 establishes a lower bound on the number of edges needed to obtain a complete ranking of the community. It also identifies the unique network architecture which reaches this lower bound: a windmill network where one of the nodes, the hub, connects all other nodes which form pairs.[2] This network architecture implies a very unequal distribution of degrees. The hub is connected to all nodes, whereas the remaining nodes have degree two or three. If agents have a limited capacity to compare other agents, the windmill network cannot be used, and one needs to resort to other more symmetric network architectures involving a larger number of links. An exact characterization of the minimal degree of a regular network for which all links can be supported remains an open question in graph theory.[3]

4 Incomplete rankings and friend-based comparisons

4.1 Comparison networks

We now consider communities where the condition of Theorem 1 fails. The condition fails either when the community is completely informative but some comparisons are based only on self-reports or because the community is not completely informative. In the latter case, there exist some type profiles $\mathbf{T}$ for which individuals collectively cannot construct a complete ranking. We let $i \bowtie_{\mathbf{T}} j$ denote the fact that i and j cannot be compared using the information contained in $\mathbf{T}$. As the mechanism ρ defines a complete ranking, it must choose an arbitrary ranking between i and j at $\mathbf{T}$. We first define a condition on the mechanism ρ guaranteeing that the arbitrary ranking chosen between i and j is independent of the reported type profile $\mathbf{T}$.

Independence: The ranking ρ satisfies independence if for any two type profiles $\mathbf{T}$ and $\mathbf{T}'$ such that $i \bowtie_{\mathbf{T}} j$ and $i \bowtie_{\mathbf{T}'} j$, then $\rho_i(\mathbf{T}) > \rho_j(\mathbf{T}) \Leftrightarrow \rho_i(\mathbf{T}') > \rho_j(\mathbf{T}')$.

When the independence condition is satisfied, if the comparison between some pair

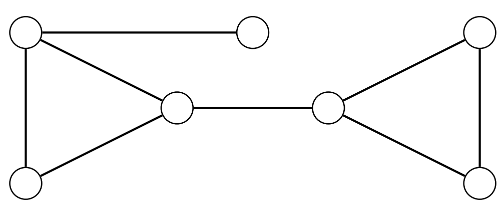
[2]The proof of the theorem is very different from known proofs of the Friendship Theorem, mostly because we focus attention on the minimization of the number of edges rather than on the construction of a graph where any intersection of neighborhoods is a singleton.

[3]A family of regular graphs, called the rook graphs, satisfy the property. For an integer $m \geq 2$, rook graphs are regular graphs of degree $2(m-1)$ among m^2 nodes, and have the property that any two connected nodes have $m-2$ nodes in common and every pair of unconnected nodes has two common neighbors. See Brouwer and Haemers (2011) for more details on rook graphs.

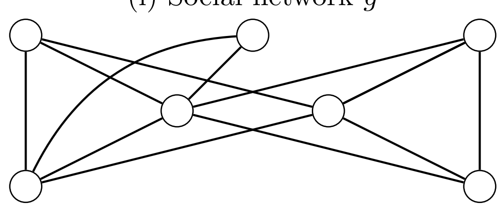
of individuals i and j relies on self-reports, the mechanism ρ cannot simultaneously satisfy ex post incentive compatibility and efficiency.

Proposition 1. *Suppose that there exists a pair of individuals (i, j) such that $g_{ij} = 1$ but there is no individual k such that $g_{ik}g_{jk} = 1$. Then there exists no mechanism satisfying independence, ex post incentive compatibility and efficiency.*

Proposition 1 extends the necessity argument of Theorem 1 to show that the planner cannot construct an ex post incentive-compatible and efficient mechanism when two individuals provide self-reports. Hence truthful comparisons based on self-reports cannot be elicited by the planner. We thus ignore comparisons based on self-reports. We now modify the type of individual i, T_i by removing any comparison t^i_{ij} which is not supported by a third individual, i.e., we let $t^i_{ij} = 0$ if there exists no $k \neq i, j$ such that $t^k_{ij} \neq 0$. We search for mechanisms satisfying independence, ex post incentive compatibility and efficiency in the community where self-comparisons are ignored. Second, we construct a *comparison network* h which captures all comparisons that can be obtained using friend-based comparisons. Formally, we let $h_{ij} = 1$ if and only if there exists k such that $g_{ik}g_{jk} = 1$. The network h collects all pairs of individuals which can be compared by a third individual. It differs from the social network g in two ways: (i) pairs of individuals which are linked in g but do not have a common friend appear in g but not in h, (ii) pairs of individuals which are not directly linked in g but have a common friend appear in h but not in g. Figure 3 illustrates a social network g and the corresponding comparison network h.



(i) Social network g



(ii) Comparison network h

Figure 3: Social and comparison networks

The comparison network h is the minimal set of comparisons that the planner can guarantee for any possible realization of the characteristics. The planner complements the comparisons contained in h by taking their transitive closure. If the conditions of Theorem 1 hold, the comparison network h is the complete network. When the conditions fail, we characterize comparison networks which can be supported by a mechanism satisfying independence, ex post incentive compatibility and efficiency.

4.2 Connected comparison networks and bipartite social networks

We first provide a characterization of social networks which generate connected comparison networks.

Proposition 2. *Suppose that $n \geq 3$. The comparison network h is connected if and only if g is not bipartite.*

Proposition 2 establishes that the network h is connected if and only if the social network g is not bipartite. If the network g is not bipartite, we construct a path connecting any pair of individuals i and j in the comparison network. If the network g is bipartite, and the nodes partitioned into the two sets A and B, the comparison network h is disconnected into two components: individuals in A rank individuals in B and individuals in B rank individuals in A. Individuals can be ranked inside the two sets A and B but rankings of individuals across the two sets must be arbitrary. Notice however that an individual in A cannot improve his ranking by lying about the ranking of individuals in B. Hence, when the social network g is bipartite (and the comparison network h disconnected), it is easy to construct a mechanism satisfying independence, ex post incentive compatibility and efficiency.

Proposition 3. *Suppose that the social network g is bipartite with two sets of nodes A and B. Then there exists a mechanism satisfying independence, ex post incentive compatibility and efficiency, which generates a ranking which coincides with the comparison network h on its two components A and B.*

Proposition 3 characterizes one situation where the planner can elicit information about comparisons: when the set of individuals in the community can be partitioned into two subsets where members of one subset observe members of the other subset.

For example, one could survey separately men and women and ask men about the characteristics of women and women about the characteristics of men. However, this design would not allow the planner to obtain information about the ranking of individuals across the two sets. The mechanism completes the partial ranking by an arbitrary ranking across individuals in the two sets, possibly resulting in a final ranking which is very different from the true ranking. We observe that the partition of the set of nodes into groups which rank each other is the basis of most algorithms proposed in the computer science literature on peer selection.

4.3 Social quilts in incomplete communities

We now consider communities for which the comparison network h is connected but not complete. We provide a sufficient condition under which the planner can construct a mechanism satisfying independence, ex post incentive compatibility and efficiency. The mechanism is an extension of the mechanism constructed in the proof of Proposition 1 for completely informative communities.

Proposition 4. *Suppose that all links in g are supported (for all i, j such that $g_{ij} = 1$, there exists a k such that $g_{ik}g_{jk} = 1$). Then there exists a mechanism satisfying independence, ex post incentive compatibility and efficiency.*

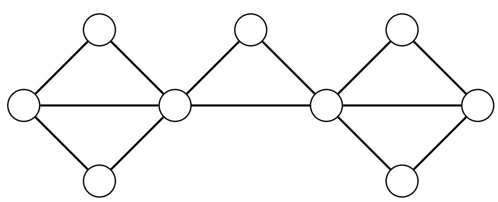


Figure 4: A supported social network g

Proposition 4 identifies social networks which allow the planner to construct an incomplete ranking of the individuals: all links must be supported and the social network is thus formed of a collection of triangles. Following Jackson, Rodriguez-Barraquer, and Tan (2012), we call these communities "social quilts". Figure 4 illustrates one of these networks. Notice that some comparisons are supported as links within the triangles, and other comparisons are supported as links across triangles. Links across traingles do not play a role in Jackson, Rodriguez-Barraquer, and Tan's (2012) favor exchange context.

Whether there exist other social networks g generating connected comparison networks h for which the planner can construct a mechanism satisfying independence, ex post incentive compatibility and ex post efficiency remains an open question. However, there are social networks for which the planner will not be able to construct a mechanism satisfying these three properties, as shown in the following example.

Example 1. Let $n = 4$. individuals i, j, k are connected in a triangle and individual l is connected to i. $\triangle$

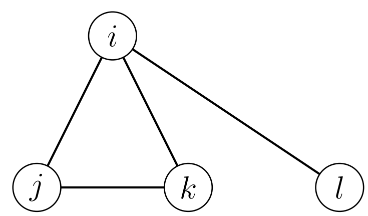


Figure 5: A social network g where a mechanism does not exist

In this example, the links $(i, j), (i, k), (j, k)$ are supported, but the link (i, l) is not supported. Consider a realization of the characteristics such that $\theta_l > \theta_j > \theta_i > \theta_k$. If individual i announces $\theta_l > \theta_j > \theta_k$, by ex post efficiency, the planner constructs the rankings k, i, j, l and the rank of individual i must be equal to 2. If on the other hand individual i announces $\theta_j > \theta_k > \theta_l$, the planner constructs the ranking l, k, i, j and the ranking of individual i is now equal to 3. Hence individual i has an incentive to lie and announce $\theta_j > \theta_k > \theta_l$.

In Example 1, the planner's ranking of i depends on his announcement on the rankings (j, l) and (k, l). Given that $\theta_j > \theta_i > \theta_k$, and that (i, j, k) form a triangle, the planner must rank i between j and k. Hence she cannot rank all three individuals j, k and l on the same side of individual i, as in the mechanism constructed in the proof of Proposition 4. But then, the announcement of individual i on (j, k, l), by changing the rank of l with respect to j and k, will also affect the ranking of individual i. Because individual i can manipulate his rank by his announcements on the unsupported links (j, k) and (j, l), there is no mechanism satisfying ex post incentive compatibility and efficiency in this community.

5 Real-life social networks

In the two previous sections, we analyzed conditions on social networks under which the planner can construct rankings that satisfy ex post incentive compatibility and efficiency. We use real social network data from India (Banerjee, Chandrasekhar, Duflo, and Jackson, 2013) and Indonesia (Alatas, Banerjee, Chandrasekhar, Hanna, and Olken, 2016) to highlight three implications of our theoretical results. We first observe that the information obtained in a social network does not depend only on the number of links. For a given density of the social network, we witness a large variation in the information obtained by friend-based ranking, depending on the exact structure of the network. Second, we analyze the role played in friend-based comparisons by supported links and links across triangles. For low-density networks, we show that links across triangles provide the majority of the friend-based comparisons. Third, we analyze how cognitive limitations affect the number of comparisons elicited by the planner. We cap the number of comparisons per individual and observe that when the cap is small relative to the community size, the information loss due to the cap is also small.

The data from India and Indonesia are particularly useful because they contain multiple independent networks: 75 villages from Karnataka, India, and 622 neighborhoods from three provinces in Indonesia. Indonesian networks are smaller and denser than the Indian networks. We focus on the giant component of each network. Table 1 provides summary statistics of the networks. We report the mean, minimum, and maximum for each measure. The combined sample of networks provides a large range in network size and structure. We measure *information* using *the density of the comparison network,* which is simply the count of unique comparisons as a share of the $\frac{n(n-1)}{2}$ possible comparisons.

Figure 7 provides further detail on the distribution of network characteristics. Notice that the Indonesian networks are denser and more clustered, and display shorter average distances than the Indian networks. As a result, the Indonesian networks contain more information, i.e., result in denser comparison networks.

Figure 7 shows a tight relationship between average distance and the quantity of information. Since every comparison (i, j) is provided by a path of length 2 between i and j, this relationship is not surprising. In contrast to the relation between information and average distance, the relationship between information and density is not tight. In the following section we use an example to highlight two reasons for the variation in the quantity of information at a given density.

5.1 Large variation in the quantity of information for a given density

Dense social networks provide many comparisons, but density is not a good proxy for the quantity of information that the planner can extract from a social network. The windmill of Theorem 3 is completely informative but its density is only $\frac{3}{n}$. This insight also applies to the data here. In the bottom left panel of Figure 8, we plot the density against the quantity of information for social networks of more than 50 households. We highlight two networks and plot their corresponding network diagrams. The orange network, from India, has 75 nodes and a density of 0.12. The green network, from Indonesia, has 69 nodes and the same density of 0.12. Despite having equal density, the orange network provides a quantity of information of 0.62—nearly double the green network with 0.33.

Two factors contribute to the greater amount of information in the orange network. First, as shown in the top right panel of Figure 8, the degree distribution of the orange is spread more widely than the green. The number of comparisons provided by a single node is a convex function of degree. For a given density, the greater the spread in the degree distribution, the more comparisons the social network provides.

Second, the green network is a combination of cliques that are weakly connected to each other. Cliques repeat comparisons. Take a clique of seven nodes as an example. These seven nodes provide 105 comparisons yet 84 of these comparisons are repeated ones. Since the green and orange networks have a similar number of nodes and equal density, they each produce a total number of comparisons similar to the other's. The difference is that a greater share of the green network's comparisons are repeated. This example shows that the success of friend-based ranking depends not only on the number of links, but, more importantly, on how those links are structured.

5.2 Decomposition of information

Proposition 4 (in Section 4) shows the importance of triangles in constructing incentive-compatible and efficient mechanisms. Both the links within and across triangles are used to obtain truthful comparisons. For a given network we can decompose information into comparisons provided within and across triangles.

We approach the decomposition by removing all unsupported links from the network and recalculating the quantity of information. The resulting supported network is incentive-compatible but information is reduced—from 0.37 to 0.27 on average for India and from 0.78 to 0.75 on average for Indonesia. A greater share of the Indonesian links is supported, which is due to the fact that those networks are denser and more clustered.

We decompose the comparisons in the supported network into those within and across triangles. If a comparison appears both within and across a triangle, we categorize the comparison as "within." Figure 9 shows the decomposition for 50 Indian and 50 Indonesian networks. Each bar corresponds to a network and the bar is split between within triangles, across triangles, and a remainder (i.e., the comparisons which appear only in the unsupported network). At lower densities the majority of comparisons are provided across triangles, while as density of the social network increases, the share of comparisons within triangles increases.

5.3 Capping comparisons

For a given degree distribution, we can define a simple upper bound on the number of bilateral comparisons. Since we measure information by counting links in the comparison network, an upper bound on information is reached when none of the bilateral comparisons produced in the social network is repeated. With the degree of individual i denoted as d_i, the upper bound is $\sum_{i=1}^{n} \frac{d_i(d_i-1)}{n(n-1}$, which is simplified to $\frac{d(d-1)}{n-1}$ for regular networks of degree d.

To analyze the effect of capping the degree of individuals on the quantity of information, we use simulations on the social network data. For each individual i, we randomly pick five friends whom i will compare to each other. The resulting network is directed. Suppose j picks i and i has more than five friends. There is no guarantee that i will also pick j. Figure 10 uses the mean from 100 iterations to measure the number of nonrepeated comparisons in the capped network. The standard deviation is less than .01 for any given network in our sample. This variation depends on the starting network.

The capped information is close to the upper bound. In Figure 10 we contrast the upper bound to the mean information provided by the capped networks. Each bar represents a network capped at degree 5. When the cap is small relative to the community size (i.e., the number of households), only a small share of the comparisons in the capped network is repeated, so that the capped information is close to the upper bound.

6 Robustness and extensions

The analysis of friend-based ranking mechanisms relies on specific assumptions on the model. In this section, we relax some of these assumptions to test the robustness of our results.

6.1 Dominant strategy implementation

We first strengthen the incentive compatibility requirement to dominant-strategy implementation. The following proposition shows that dominant-strategy implementation is too strong in our setting. The outcome set is not rich enough to permit the construction of strategy-proof mechanisms. We first recall the definition of strategy-proofness:

Strategy-proofness. For any individual i, for any vector of announcements $(\hat{T}^{-i})$ and any types T^i, T'^i,

$$\rho_i(T^i, \hat{T}^{-i}) \geq \rho_i(T'^i, \hat{T}^{-i}).$$

Proposition 5. *Let g be a triangle. There exists no mechanism satisfying strategy-proofness and ex post efficiency.*

Proposition 5 is an impossibility result, highlighting a conflict between strategy-proofness and efficiency in a very simple network architecture. As shown in the proof, the impossibility stems from the coarseness of the outcome space, which limits the power of the planner. There are only three possible outcomes corresponding to the three possible ranks. Strategy-proofness imposes a large number of constraints on the mechanism. We show, using a combinatorial argument, that if all the constraints are satisfied, two individuals must be occupying the same rank for some vectors of announcement. Hence, it is impossible to elicit truthful information in a complete network with three individuals.[4]

[4]The extension of this impossibility result to more than three individuals remains an open ques-

6.2 Coarse rankings

We next relax the assumption that the planner chooses a complete ranking and that individuals have strict preferences over ranks. We consider a setting where the planner selects only broad indifference classes. This is the typical situation in which the planner selects a set of recipients of the benefits of social programs, or of research funds. If the planner only chooses broad categories, she might be able to construct ex post incentive-compatible and efficient mechanisms even if self-reports are not supported by a third individual. The intuition is immediate: if there exist two "worst spots" in the ranking, the planner can punish individuals who send conflicting self-reports by placing both of them on the worst spot. We formalize this intuition in the following proposition.

Proposition 6. *There exists an ex post incentive-compatible and efficient mechanism in any completely informative community if and only if the planner can place two individuals in the worst spot, i.e., if and only if any individual i is indifferent between $\rho(i) = 1$ and $\rho(i) = 2$.*

Proposition 6 thus shows that it is easier to construct ex post incentive-compatible and efficient mechanisms when the planner does not construct a complete ranking of the individuals. This observation raises new possibilities. It may be possible to construct strategy-proof and efficient mechanisms when the planner only assigns agents to broad categories.

6.3 Group incentive compatibility

We now allow for individuals to jointly deviate from truth-telling. We let individuals coordinate their reports and jointly misreport their types.

Consider a triangle with three individuals. Each individual reports on the three links. The mechanism that we constructed in Theorem 1 of Section 3 assigns a ranking $\rho(i) > \rho(j)$ when at least two of the individuals report that i is higher than j. This creates an incentive for any pair of individuals to misrepresent their types. For example, if the true ranking is $\theta_3 > \theta_2 > \theta_1$, individuals 1 and 2 have an incentive to misreport and announce that 2 is higher than 1, and 1 is higher than 3, so that in the end, $\rho(2) = 3 > 2$ and $\rho(1) = 2 > 1$.

tion.

This intuition can be exploited to show that there does not exist any mechanism satisfying ex post group-incentive compatibility and efficiency when $n = 3$. We first provide a formal definition of ex post group-incentive compatibility:

Ex post group incentive compatibility. For any vector of types $\mathbf{T}$, there does not exist a coalition S and a vector of types $\mathbf{T'^S}$ such that for all individuals i in S,

$$\rho_i(\mathbf{T'^S}, \mathbf{T^S}) \geq \rho_i(\mathbf{T})$$

and

$$\rho_i(\mathbf{T'^S}, \mathbf{T^S}) > \rho_i(\mathbf{T}).$$

for some $i \in S$.

Proposition 7. *Let g be a triangle. There does not exist a mechanism satisfying ex post group-incentive compatibility and efficiency.*

6.4 Homophily

In this last extension, we analyze the effect of homophily on friend-based ranking. Using Golub and Jackson's (2012) *islands model* of network formation, we show that the probability of finding the full ranking of characteristics initially increases and then decreases in the homophily parameter.

In a community of n individuals, $n - 1$ comparisons are necessary and sufficient to determine the full ranking of characteristics. The lowest is compared to the second lowest, the second to the third, and so on. We call this subset of comparisons C. We will consider a random model of network formation and ask: how likely is it that individuals form links that generate the set of bilateral comparisons C? Suppose that pairs of individuals form friendships with a given probability p (which is the Erdős-Rényi random graph model) and through friend-based ranking the realized network provides a set of comparisons T. What is the probability $\Pr[C \in T]$?

For simplicity, we order the individuals in the community $\{1, 2, ..., n - 1, n\}$ so that the private characteristic $\theta_i > \theta_j$ if and only if $i > j$. We then define $C = \{(1, 2), (2, 3), ..., (n - 2, n - 1), (n - 1, n)\}$. Consider a pair (i, j). $\Pr[(i, j) \in T] = \Pr[\exists\ k \neq i, j : g_{ik} = g_{jk} = 1] = 1 - (1 - p^2)^{n-2}$. Since there are $n - 1$ pairs in C, $\Pr[C \in T] = \left(1 - (1 - p^2)^{n-2}\right)^{n-1}$.

We now divide the community into two groups of equal size, N^L for low and N^H for high such that $\theta_l < \theta_h \; \forall \; l \in N^L, h \in N^H$. Individuals form friendships within their group with probability p_w and outside of their group with probability p_o (which is Golub and Jackson's (2012) *islands model* with 2 groups). For $p_w \geq p_o$, the gap between p_w and p_o is a measure of homophily.

Keeping p_o constant, if we increase p_w, $\Pr[C \in T]$ will increase since we have raised the expected density of the social network. We need to keep the expected density of the network constant to isolate the impact of homophily. Since we disregard self-comparisons by Proposition 1, there are more outside group links than within group links. The number of outside group links is $(\frac{n}{2})^2$ while the number of within group links is $(\frac{n}{2})^2 - \frac{n}{2}$. The ratio simplifies to $n-2$ within group links for every n outside group links.

Starting from a zero homophily base of $p = p_w = p_o$, we can analyze the impact of homophily by increasing p_w and decreasing p_o to keep the expected number of links constant. Let $p_w = p + \eta$, where η is the homophily parameter. To keep the number of links constant, $p_o = p - \eta\frac{n}{n-2}$.

From a base of $\eta = 0$ we can increase η and observe how $\Pr[C \in T]$ responds. This is represented graphically in Figure 6 for a community size $n = 250$ and a base probability of friendship $p = 0.15$. Along the horizontal axis, as η increases from 0 to around 0.1 the probability that the realized comparisons contain the comparisons needed to derive the full ranking increases.

The intuition is simple. As homophily increases the probability that some individual k is friends with two individuals in the same group increases while the probability that k is friends with two individuals in different groups decreases. Since nearly all of the comparisons in C are pairs within the same group, $\Pr[C \in T]$ rises with homophily. However, when the probability of outside group friendships p_o approaches zero, $\Pr[C \in T]$ approaches zero since there is little chance that the comparison of the highest in N^L to the lowest in N^H is in T. In Figure 6, $\Pr[C \in T]$ drops sharply as η is above 0.12 and p_o approaches zero. Low levels of homophily improve friend-based ranking whereas extreme homophily reduces the power of friend-based ranking as agents across groups cannot be compared.

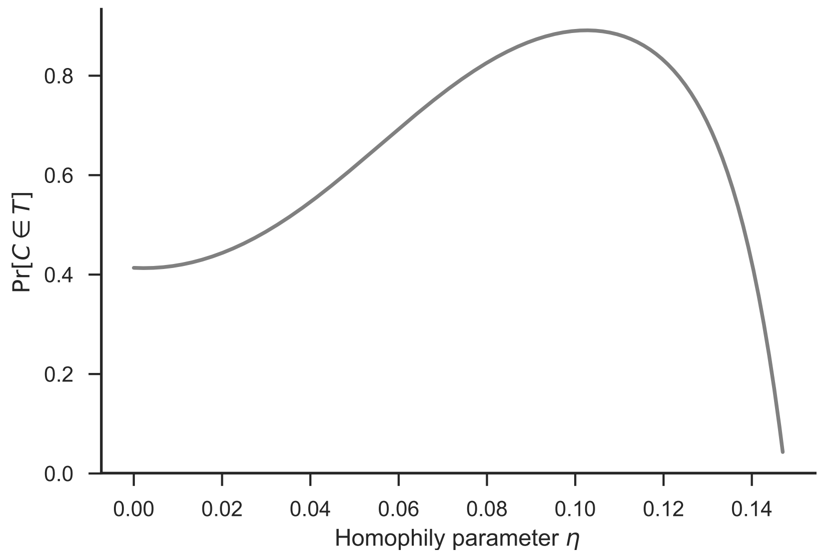


Figure 6: Impact of homophily ($p = 0.15$, $n = 250$)

7 Conclusion

This paper analyzes the design of mechanisms to rank individuals in communities in which individuals have only local, ordinal information on the characteristics of their neighbors. In these communities, pooling the information of all individuals may not be sufficient to obtain a complete ranking, and so we distinguish between completely informative communities and communities where only incomplete social rankings can be obtained.

In completely informative communities, we show that the planner can construct an ex post incentive-compatible and ex post efficient mechanism if and only if each pair of individuals is observed by a third individual, i.e., the individuals in each pair have a common friend. We use this insight to characterize the sparsest social network for which a complete ranking exists as constituting a "friendship network" (or "windmill network") in the sense of Erdős, Rényi, and Sós (1966).

When the social network is not completely informative, we show that any self-report which is not supported by a third party must be discarded. We provide two sufficient conditions on the social network under which an ex post incentive-compatible and ex post efficient mechanism may be constructed.

First, in bipartite networks, individuals on one side of the network can be used to rank individuals on the other side, resulting in an ex post efficient but incomplete ranking. Second, in "social quilts," where all links are supported in triangles, the planner can use the congruence of reports to construct truthful rankings over any pair of individuals.

We use data on social networks from India and Indonesia to illustrate the results of the theoretical analysis. We measure information provided by the social network as the share of unique comparisons which can be obtained by friend-based comparisons (which corresponds to the density of the comparison network) and show that (i) information varies greatly even for a given density, (ii) across-triangle comparisons are important at low densities, and (iii) information is close to an upper bound when the degree is capped at a small value relative to the community size.

Finally, we discuss robustness and extensions of the model, focusing on strategy-proofness, group-incentive compatibility, coarse rankings, and homophily.

To the best of our knowledge, this is the first paper to analyze this intriguing theoretical problem—the design of a mechanism constructing a complete ranking when individuals have local, ordinal information based on a social network. In future work, we would like to further our understanding of the problem, by considering in more detail the difference between ordinal and cardinal information, between complete and coarse rankings, and between different concepts of implementation. We also plan to extend the empirical and policy implications of the theoretical model by analyzing specific institutional settings in more detail.

References

Alatas, Vivi, Abhijit Banerjee, Arun G Chandrasekhar, Rema Hanna, and Benjamin A Olken. 2016. "Network structure and the aggregation of information: Theory and evidence from Indonesia." *The American Economic Review* 106 (7):1663–1704.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias. 2012. "Targeting the poor: evidence from a field experiment in Indonesia." *American Economic Review* 102 (4):1206–40.

Alon, Noga, Felix Fischer, Ariel Procaccia, and Moshe Tennenholtz. 2011. "Sum of us:

Strategyproof selection from the selectors." In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*. ACM, 101–110.

Aziz, Haris, Omer Lev, Nicholas Mattei, Jeffrey S Rosenschein, and Toby Walsh. 2016. "Strategyproof Peer Selection: Mechanisms, Analyses, and Experiments." In *AAAI*. 397–403.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The diffusion of microfinance." *Science* 341 (6144):1236498.

Baumann, Leonie. 2017. "Identifying the Best Agent in a Network." Tech. rep., Working paper.

Brouwer, Andries E and Willem H Haemers. 2011. *Spectra of graphs*. Springer Science & Business Media.

Brown, Caitlin, Martin Ravallion, and Dominique Van de Walle. 2016. "A poor means test? Econometric targeting in Africa." Tech. rep., National Bureau of Economic Research.

Coady, David, Margaret Grosh, and John Hoddinott. 2004. "Targeting outcomes redux." *The World Bank Research Observer* 19 (1):61–85.

Crémer, Jacques and Richard McLean. 1985. "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent." *Econometrica* 53 (2):345–61.

De Clippel, Geoffroy, Herve Moulin, and Nicolaus Tideman. 2008. "Impartial division of a dollar." *Journal of Economic Theory* 139 (1):176–191.

Erdős, Paul, Alfréd Rényi, and Vera Sós. 1966. "On a Problem of Graph Theory." *Studia Scientiarium Mathematicarum Hungarica* 1 (1):215–235.

Golub, Benjamin and Matthew O Jackson. 2012. "How homophily affects the speed of learning and best-response dynamics." *The Quarterly Journal of Economics* 127 (3):1287–1338.

Holzman, Ron and Hervé Moulin. 2013. "Impartial nominations for a prize." *Econometrica* 81 (1):173–196.

Huneke, Craig. 2002. “The friendship theorem.” *The American mathematical monthly* 109 (2):192–194.

Hussam, Reshmaan, Natalia Rigol, and B Roth. 2017. “Targeting high ability entrepreneurs using community information: Mechanism design in the field.” Tech. rep., Working paper.

Jackson, Matthew O, Tomas Rodriguez-Barraquer, and Xu Tan. 2012. “Social capital and social quilts: Network patterns of favor exchange.” *American Economic Review* 102 (5):1857–97.

Kurokawa, David, Omer Lev, Jamie Morgenstern, and Ariel D Procaccia. 2015. “Impartial Peer Review.” In *IJCAI*. 582–588.

Longyear, Judith Q and Torrence D Parsons. 1972. “The friendship theorem.” In *Indagationes Mathematicae (Proceedings)*, vol. 75. Elsevier, 257–262.

Rai, Ashok S. 2002. “Targeting the poor using community information.” *Journal of Development Economics* 69 (1):71–83.

Wilf, Herbert S. 1971. “The friendship theorem.” In *Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969)*. 307–309.

A Tables

Table 1: Summary statistics of social networks

	India		Indonesia	
Networks	75		622	
Number of households	198.72	[77, 356]	52.85	[11, 263]
Share in giant component	.95	[.85, .99]	.65	[.22, 1.00]
Average degree	9.34	[6.82, 13.83]	17.96	[2.00, 218.00]
Density	.05	[.02, .12]	.53	[.10, 1.00]
Average clustering	.26	[.16, .45]	.82	[.48, 1.00]
Average distance	2.75	[2.30, 3.32]	1.77	[1.00, 4.32]
Information	.37	[.18, .62]	.78	[.25, 1.00]

Notes: Means are reported with minimum and maximum in brackets. Information is measured by the density of the comparison network. All statistics (except the number of households) are calculated on the giant component. Data is sourced from Banerjee et al. (2013) for India and Alatas et al. (2016) for Indonesia.

B Figures

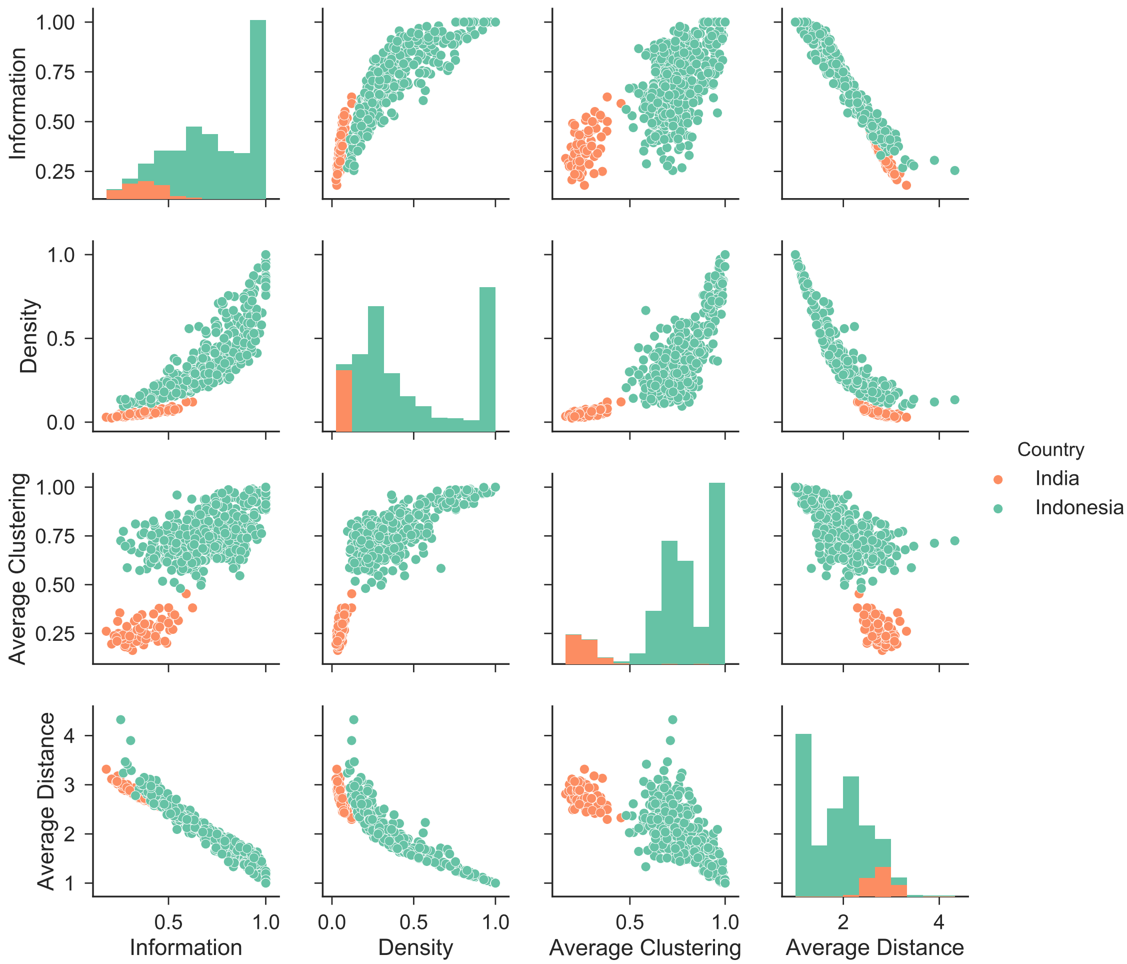


Figure 7: Distribution of social network measures

Note: Social networks from India (in orange) and Indonesia (in green). Information is measured as the density of the comparison network.

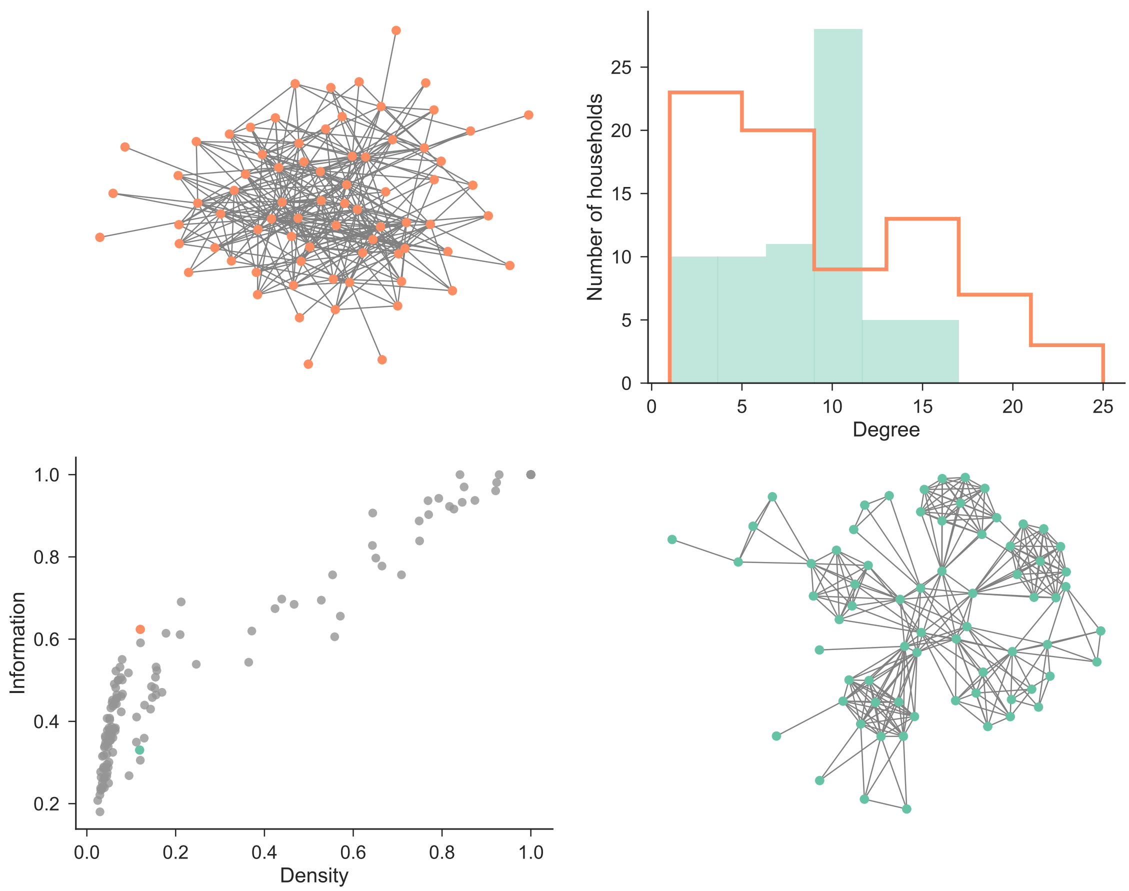
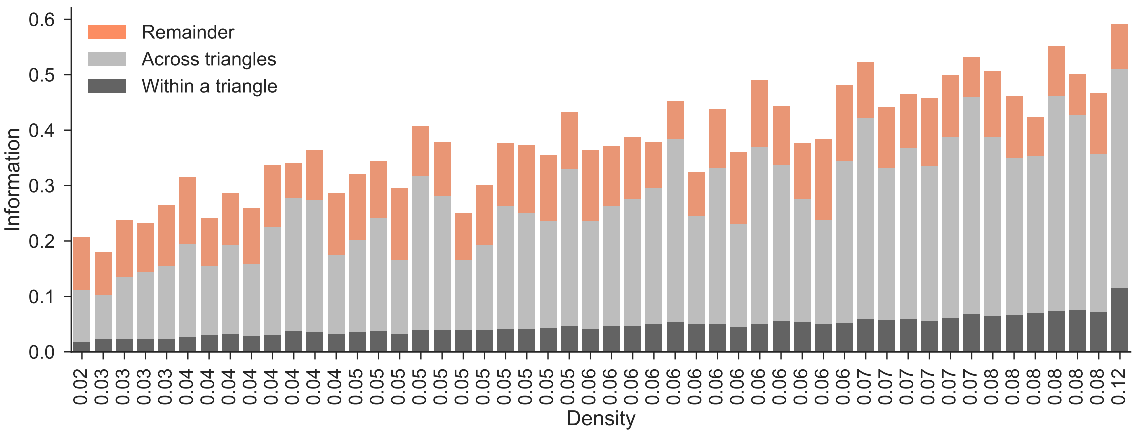
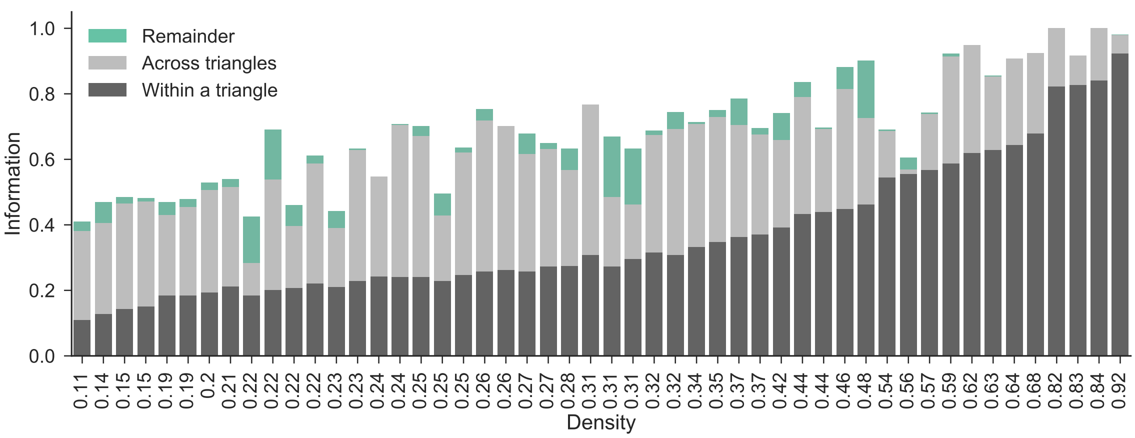


Figure 8: Large variation of information for a given density

Note: The bottom left panel shows a scatter-plot of information and density for networks of more than 50 households. Two networks of similar density are highlighted by orange and green points on the scatter plots. The network diagrams corresponding to those two points are plotted in the top left and bottom right panel. The degree distribution of the highlighted networks is shown in the top right panel.



(i) India



(ii) Indonesia

Figure 9: Decompose information

Note: We decompose information (unique comparisons) into comparisons which are provided within triangles, across triangles, and a remainder. By Proposition 4 all within and across triangle comparisons are incentive-compatible. Comparisons within the remainder may not be incentive-compatible.

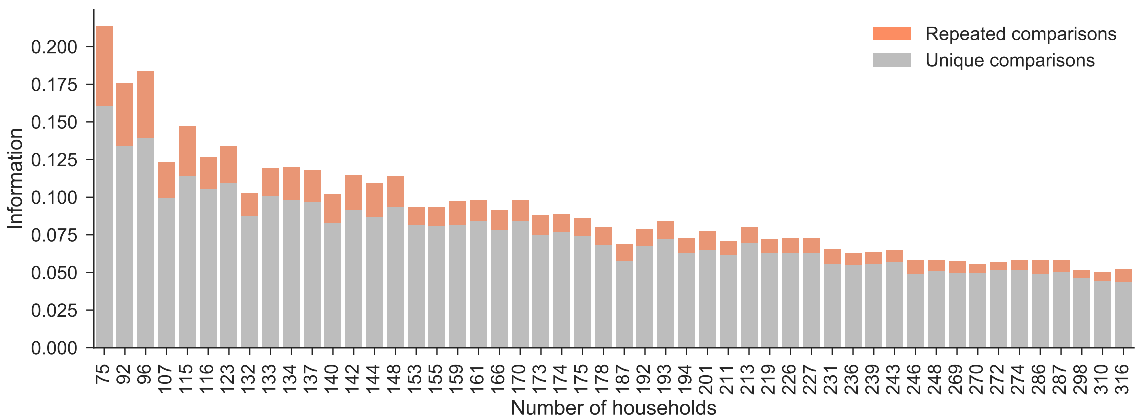


Figure 10: Capping comparisons at 5 friends

Note: Each bar represents a single network on which we simulate a cap on degree. Each individual provides comparisons of at most 5 friends, selected uniformly at random when the cap is binding. An upper bound equal to sum of unique and repeated comparisons is given as a function of the degree distribution. The split between unique and repeated comparisons is calculated as the mean after 100 iterations.

C Proofs

Proof of Lemma 1

The condition is obviously sufficient, as it guarantees that for any pair (i, j) there exists an individual k such that $t_{ij}^k \neq 0$, Hence the matrix generated by $(T^1, .., T^n)$ contains nonzero entries everywhere outside the diagonal. Conversely, suppose that there exists a pair of individuals (i, j) who is observed by no other player and such that $g_{ij} = 0$. Consider a realization of the characteristics such that θ_i and θ_j are two consecutive values. No individual can directly compare i and j. In addition, because there is no k such that $\theta_k \in (\theta_i, \theta_j)$, there is no k such that $\theta_i \prec \theta_k \prec \theta_j$ or $\theta_j \prec \theta_k \prec \theta_i$. Hence the social network g is not completely informative.

Proof of Theorem 1

Sufficiency. Suppose that for any pair of individuals (i, j), there exists a third individual k for whom $g_{ik} = g_{jk} = 1$. We define the mechanism ρ by constructing comparisons. Let r_{ij} denote the comparison between i and j chosen by the planner.

First consider a pair of individuals (i, j) who observe each other, $g_{ij} = 1$. By assumption, there are at least three reports on the ranking of i and j. If all individuals transmit the same report t_{ij}, let $r_{ij} = t_{ij}$. If all individuals but one transmit the same report t_{ij} and one individual reports $t'_{ij} = -t_{ij}$, ignore the ranking t'_{ij} and let $r_{ij} = t_{ij}$. In all other cases, let $r_{ij} = 1$ if and only if $i > j$.

Second consider a pair of individuals (i, j) who do not observe each other, $g_{ij} = 0$. By assumption, there exists at least one individual k who observes them both. If there are at least three individuals who observe i and j, use as above a mechanism such that $r_{ij} = t_{ij}$ if all individuals agree on t_{ij} or only one individual chooses $t'_{ij} = -t_{ij}$, and let $r_{ij} = 1$ if $i > j$ otherwise.

If one or two individuals observe i, j, pick the individual k with the highest index. Consider the vector of announcements $\tilde{T}^{-k}$ where one disregards the announcements of individual k. Let $\succ_{\tilde{T}^{-k}}$ be the binary relation created by letting $t_{ij} = 1$ if and only if $t_{ij}^l = 1$ for all $l \neq k$. If there exists a directed path of length greater or equal to 2 between i and j in $\succ_{\tilde{T}^{-k}}$, and for all directed paths between i and j in $\succ_{\tilde{T}^{-k}}$, $i^0, .., i^L$ we have $t_{i^l i^{l+1}} = 1$, then $r_{ij} = 1$. If on the other hand for all directed paths between i and j in $\succ_{\tilde{T}^{-k}}$, $t_{i^l i^{l+1}} = -1$, then $r_{ij} = -1$. In all other cases, let individual k dictate the comparison between i and j, $r_{ij} = t_{ij}^k$.

Now consider all comparisons r_{ij}. If they induce a transitive binary relation on

N, let ρ be the complete order generated by the comparisons. Otherwise, consider all shortest cycles generated by the binary relation r_{ij}. If there exists a single individual i who dictates at least two comparisons in all shortest cycles, individual i is punished by setting $\rho_i = 1$ and $\rho_j > \rho_k$ if and only if $j > k$ for all $j, k \neq i$. If this is not the case, pick the arbitrary ranking where $\rho_i > \rho_j$ if and only if $i > j$.

We now show that the mechanism ρ is ex post incentive-compatible and ex post efficient.

Suppose that all individuals except k report their true type, and consider individual k's incentive to report $T'^k \neq T^k$. On any link (i, j) such that $g_{ij} = 1$, as all other individuals make the same announcement, individual k cannot change the comparison r_{ij} by misreporting.

Consider a link (i, j) such that $g_{ij} = 0$ and $g_{ik}g_{jk} = 1$. If there are at least three individuals who observe i and j, individual k cannot affect the outcome. Otherwise, if there is a directed path of length greater than equal to 2 in $\succ_{\tilde{T}^{-k}}$, individual k's report cannot change the ranking. If individual k is not the highest ranked individual who observes i and j, then she cannot change the comparison r_{ij} by misreporting. Hence we only need to focus attention on pairs (i, j) such that k is the highest index individual who observes i and j and there is no directed path between i and j in $\succ_{\tilde{T}^{-k}}$.

Suppose that all individuals $l \neq k$ announce the truth, so that $\succ_{\tilde{T}^{-k}} = \succ_{T^{-k}}$. We first show that individual k cannot gain by making an announcement which induces cycles in the ranking generated by the comparisons r_{ij}. Suppose that the ranking generated by r_{ij} exhibits cycles. We first claim that the shortest cycles must be of length 3.

Suppose that there exists a cycle of length L, $i^0, i^1, .., i^L$. Because the community is completely informative, the binary comparisons generated by the announcements are complete, so that for any l, m, either $r_{i^l i^m} = 1$ or $r_{i^l i^m} = -1$. Now consider i^0, i^1, i^2. If $r_{i^0 i^2} = -1$, $i^0, i^1 i^2 i^0$ forms a shortest cycle of length 3. If not, $r_{i^0 i^2} = 1$ and we can construct a cycle of length $L - 1$, $i^0 i^2, .., i^L$. By repeating this argument, we either find shortest cycles of length 3 or end up reducing the initial cycle to a cycle of length 3.

Consider next a shortest cycle of $ijli$. We claim that individual k must dictate at least two comparisons in the cycle.

First note that if k does not dictate the comparison between i and j, there must be a directed path between i and j in $\succ_{T^{-k}}$. To see this, notice that either $g_{ij} = 1$ and then $i \succ_{T^{-k}} j$ or $g_{ij} = 0$ but i and j are not observed by k or are observed by k and another individual with a higher index than k, in which case $i \succ_{T^{-k}} j$. Finally, it could be that $g_{ij} = 0$, i and j are observed by k, k is the highest index individual observing i and j, but then as k does not dictate the comparison (i, j), there must exist a directed path of length 2 between i and j in $\succ_{T^{-k}}$.

Now suppose first that k does not dictate any comparison in the cycle. There must exist a directed path between i and j, j and l and l and i in $\succ_{T^{-k}}$, a contradiction since, as all individuals tell the truth, the binary relation generated by $\succ_{T^{-k}}$ is transitive.

Next suppose that k dictates a single comparison (i, j) in the cycle but not the comparisons (j, l) and (l, i). Then there exists a directed path between j and l and a directed path between l and i in $\succ_{T^{-k}}$. Hence there exists a directed path of length greater than or equal to 2 between j and i in $\succ_{T^{-k}}$. Furthermore, as all individuals tell the truth, for all directed paths between j and i, $r_{ji} = 1$. Hence the mechanism cannot let individual k dictate the choice between i and j, yielding a contradiction.

We conclude that all shortest cycles are of length 3, and that in any cycle of length 3, individual k must dictate at least two of the three comparisons. Hence the mechanism assigns $\rho(k) = 1$ and individual k cannot benefit from inducing a cycle.

Finally suppose that all comparisons r_{ij} result in a transitive relation so that ρ can be constructed as the complete order generated by these comparisons. We claim that the comparisons generated by T^{-k} are sufficient to compute the rank of k. In fact, for any $i \neq k$, either $g_{ik} = 1$ and as all other individual tell the truth, r_{ik} is independent of the report t^k_{ik}, or $g_{ik} = 0$ and the report on (i, k) is truthfully made by another individual l. In both cases, the information contained in T^{-k} is sufficient to construct the comparison r_{ik}. Hence ρ_k is independent of the announcement T^k, concluding the proof that the mechanism is ex post incentive-compatible.

To show that the mechanism is ex post efficient notice that, when all individuals truthfully report their types, the rankings r_{ij} induce a transitive relation, and yield the complete ranking generated by $\succ_{\mathbf{T}}$.

Necessity. Suppose that the social network g satisfies the conditions of Lemma 1 but that there exists a pair of individuals (i, j) who observe each other but are not observed by any third individual k. Consider a realization of the characteristics such that θ_i and θ_j are the two lowest characteristics. Let $\mathbf{T}_1$ be the type profile if $\theta_i < \theta_j$ and $\mathbf{T}_2$ the type profile if $\theta_j < \theta_i$.

By ex post efficiency, because the rankings generated by $\mathbf{T}_1$ and $\mathbf{T}_2$ are complete,

$$\rho_i(\mathbf{T}_1) = \rho_j(\mathbf{T}_2) = 1,$$
$$\rho_i(\mathbf{T}_2) = \rho_j(\mathbf{T}_1) = 2.$$

Because there are only two announcements t^i_{ij} and t^j_{ij} on the link (i, j), ex post incentive compatibility requires that individuals i and j cannot improve their ranking by changing their reports on the link (i, j). Let T_{-ij} denote the announcements on all links but link ij. We must have

$$\rho_i(T_{-ij}, t^i_{ij} = 1, t^j_{ij} = -1) = \rho_i(\mathbf{T}_1) = 1,$$
$$\rho_j(T_{-ij}, t^i_{ij} = 1, t^j_{ij} = -1) = \rho_j(\mathbf{T}_2) = 1.$$

resulting in a contradiction as i and j cannot both be ranked at position 1.

Proof of Theorem 3

We establish the Theorem through a sequence of claims. Let $\ell(g)$ be the number of links in the social network g.

Claim 1. *If the social network is completely informative, then every individual must have at least 2 friends.*

Proof. Let d_i be the number of friends of individual i. As g is connected, $d_i \geq 1$ for all $i \in N$. Suppose that $d_i = 1$, and consider the unique neighbor j of i. As $d_i = 1$, there is no $k \neq j$ which is connected to i and can draw a comparison between i and j. Hence the network g is not completely informative, establishing a contradiction.

Claim 2. *If for any (i, j) there exists k such that $g_{ik}g_{jk} = 1$, then $\ell(g) \geq \frac{3(n-1)}{2}$ if n is odd and $\ell(g) \geq \frac{3n}{2} - 1$ if n is even.*

Proof. Consider the following problem: For a fixed number of links L, compute the maximal number of comparisons of neighbors that can be generated by a social network g when all nodes have degree contained in $[2, n-1]$. More precisely, let $(d_1, .., d_n)$ denote the degree sequence of g with the understanding that $d_{i-1} \geq d_i$ for all $i = 1, .., n$. Then consider the problem:

$$\max_{(d_1,...,d_n)} \frac{d_1(d_1-1)}{2} + \frac{d_2(d_2-1)}{2} + ... + \frac{d_n(d_n-1)}{2}$$

$$\text{subject to } 2 \leq d_i \leq n-1 \quad \forall i \ ,$$

$$d_1 + d_2 + ... + d_n = 2L \ .$$

Notice that the objective function $V(d_1, .., d_n) = \frac{d_1(d_1-1)}{2} + \frac{d_2(d_2-1)}{2} + ... + \frac{d_n(d_n-1)}{2}$ is strictly increasing and convex in $(d_1, ..., d_n)$.

Assume first that n is odd. Then pick $L = \frac{3(n-1)}{2}$ and $d_1 = n-1, d_2 = .. = d_n = 2$. Because V is strictly convex,

$$\begin{aligned} V(n-1, 2, ..., 2) &= \frac{(n-1)(n-2)}{2} + n - 1 \\ &= \frac{n(n-1)}{2} \\ &> V(d_1, ..., d_n) \end{aligned}$$

for any $(d_1, .., d_n) \neq (n-1, 2, ...2)$ such that $d_1 + .. + d_n = 3(n-1)$ and $d_i \geq 2$ for all i. Now $\frac{n(n-1)}{2}$ is the total number of comparisons. So, as $V(d_1, .., d_n)$ is strictly increasing in n, the social network g must contain at least $\frac{3(n-1)}{2}$ links for all comparisons to be constructed.

Assume next that n is even. Pick $L = \frac{3n}{2} - 1$ and $d_1 = n-1, d_2 = 3, d_3 = .. = d_n = 2$. Because V is strictly convex,

$$\begin{aligned} V(n-1, 3, 2, ..., 2) &= \frac{(n-1)(n-2)}{2} + 3 + n - 2 \\ &= \frac{n(n-1)}{2} + 1 \\ &> V(d_1, ..., d_n) \end{aligned}$$

for any $(d_1, .., d_n) \neq (n-1, 3, 2, ...2)$ such that $d_1 + .. + d_n = 3n - 2$ and $d_i \geq 2$ for all i.

In addition notice that for $L' = \frac{3n}{2} - 2$,

$$V(n-2,2,2,...,2) = \frac{(n-2)(n-3)}{2} + n - 1$$
$$= \frac{(n-2)^2 - n}{2}$$
$$> V(d_1,...,d_n)$$

for any $(d_1,..,d_n) \neq (n-2,2,2,...2)$ such that $d_1 + .. + d_n = 3n-4$ and $d_i \geq 2$ for all i.

Hence, the maximum of V_i is smaller than $\frac{n(n-1)}{2}$ when $\ell(g) = \frac{3n}{2} - 2$ and greater than $\frac{n(n-1)}{2}$ when $\ell(g) = \frac{3n}{2} - 1$, establishing that the social network g must contain at least $\frac{3n}{2} - 1$ links for all comparisons to be constructed.

Next we observe that the friendship network and the modified windmill network generate all comparisons.

Claim 3. *If n is odd, the friendship network containing exactly $\frac{3(n-1)}{2}$ links, generates all comparisons. If n is even, the windmill with sails of size 2 and one sail of size 3 with an additional link, containing exactly $\frac{3n}{2} - 1$ links, generates all comparisons.*

Proof. The hub of the network, node n_h, provides the comparisons between all other $(n-1)$ nodes. If n is odd, in any petal (i,j), i provides the comparison between j and n_h and j provides the comparison between i and n_h. If n is even, in any sail of size 2, (i,j),i provides the comparison between j and n_h and j provides the comparison between i and n_h. In the unique sail of size 3, (i,j,k), i provides the comparison between j and n_h, j provides the comparisons between i and n_h and k and n_h and k provides a (redundant) comparison between j and n_h.

Finally we establish that the friendship network and the modified windmill network are the only network architectures generating all comparisons with the minimal number of edges.

Claim 4. *If n is odd, the friendship network is the only network with degree sequence $(n-1,2,...,2)$. If n is even, the modified windmill network with $\frac{n}{2} - 2$ sails of size 2 and one sail of size 3 with an additional link is the only network with degree sequence $(n-1,3,2,..2)$.*

Proof. Let n be odd. Because one node has degree $n-1$, the network is connected and this node is a hub. All other nodes must be connected to the hub, and if they have degree 2, they must be mutually connected to one other node. Let n be even. The same argument shows that all nodes with degree 2 must be connected to the hub and one other node. These nodes are mutually connected except for the petal of size 3, where one node is connected to the two other nodes in the sail.

Proof of Proposition 1

As in the necessity part of the proof of Proposition 1, consider a realization of the characteristics such that θ_i and θ_j are the two lowest characteristics. Fix two type profiles $\mathbf{T}$ and $\mathbf{T}'$ which agree on all comparisons except that $\theta_i < \theta_j$ in $\mathbf{T}$ and $\theta_j < \theta_i$ in $\mathbf{T}'$. Clearly, for any $k \neq i, j$, if $k \succ_{\mathbf{T}} j$ then $k \succ_{\mathbf{T}} i$, as i and j can be compared under $\mathbf{T}$. Similarly, if $k \succ_{\mathbf{T}'} i$ then $k \succ_{\mathbf{T}'} j$. Furthermore, as θ_i and θ_j are the two smallest characteristics, all individuals k which can be compared to i and j have higher rank than i and j. Next consider $k \neq i, j$ such that $k \bowtie_{\mathbf{T}} i$. Then we must also have $k \bowtie_{\mathbf{T}} j$, as otherwise $k \succ_{\mathbf{T}} j$ which implies $k \succ_{\mathbf{T}} i$. Similarly, if $k \bowtie_{\mathbf{T}'} j$ then $k \bowtie_{\mathbf{T}'} i$. Hence if an individual k cannot be compared to i under $\mathbf{T}$, it cannot be compared to j under $\mathbf{T}$, nor to j under $\mathbf{T}'$ nor to i under $\mathbf{T}'$.

By ex post efficiency, for all k which can be compared to i, j, $\rho_k(\mathbf{T}) > \rho_j(\mathbf{T}) > \rho_i(\mathbf{T})$. Similarly, by ex post efficiency, for all k which can be compared to i, j, $\rho_k(\mathbf{T}') > \rho_i(\mathbf{T}) > \rho_j(\mathbf{T})$. By independence, for all k which cannot be compared to i, j, $\rho_k(\mathbf{T}) > \rho_i(\mathbf{T})$ if and only if $\rho_k(\mathbf{T}') > \rho_i(\mathbf{T}')$ and $\rho_k(\mathbf{T}) > \rho_j(\mathbf{T})$ if and only if $\rho_k(\mathbf{T}') > \rho_j(\mathbf{T}')$. Hence the set of individuals who are incomparable to i, j and are ranked below i and j under $\mathbf{T}$ and $\mathbf{T}'$ are identical. But this implies that

$$
\begin{aligned}
\rho_i(\mathbf{T}) &= \rho_i(\mathbf{T}') - 1, \\
\rho_j(\mathbf{T}') &= \rho_j(\mathbf{T}) - 1
\end{aligned}
$$

As we also have $\rho_i(\mathbf{T}) < \rho_j(\mathbf{T})$ and $\rho_j(\mathbf{T}') < \rho_i(\mathbf{T}')$, we must have

$$\rho_i(\mathbf{T} + 2 > \rho_j(\mathbf{T}) > \rho_i(\mathbf{T}),$$

so that $\rho_j(\mathbf{T}) = \rho_i(\mathbf{T}) + 1$. Hence

$$\rho_i(\mathbf{T}) = \rho_j(\mathbf{T}'),$$
$$\rho_j(\mathbf{T}) = \rho_i(\mathbf{T}').$$

Because there are only two announcements t^i_{ij} and t^j_{ij} on the link (i, j), ex post incentive compatibility requires that individuals i and j cannot improve their ranking by changing their reports on the link (i, j). Let T_{-ij} denote the announcements on all links but link ij. We must have

$$\rho_i(T_{-ij}, t^i_{ij} = 1, t^j_{ij} = -1) = \rho_i(\mathbf{T})$$
$$\rho_j(T_{-ij}, t^i_{ij} = 1, t^j_{ij} = -1) = \rho_j(\mathbf{T}'),$$

resulting in a contradiction as i and j cannot both be ranked at the same position.

Proof of Proposition 2

We first prove the following Claim.

Claim 5. *The comparison network is connected if and only if for all $i, j \in N$, there exists an even walk between i and j.*

Proof. Suppose first that h is connected. Pick any two nodes $i, j \in N$ and a walk $i = i^0, ..., i^m = j$ in h. By definition, for any (i^k, i^{k+1}) in the walk, there exists $j^k \in N$ such that $i^k, i^{k+1} \in N_{j^k}$. But this implies that there exists a walk in g connecting i to j given by $i^0, j^0, i^1, j^1, ..., i^{m-1}, j^{m-1}, i^m$.[5] This walk contains an even number of edges, proving necessity of the claim.

Next suppose that h is not connected and let i and j be two nodes in different components of h. We want to show that all walks between i and j in g are odd. Consider first a path between i and j. If the path is even, there exists a sequence of nodes $i = i^0, i^1, .., i^m = j$ where $m = 2l$ is even such that $g_{i^k, i^{k+1}} = 1$ for all k. But then, for any $l = 0, \frac{m}{2} - 1$, $h_{i^{2l}, i^{2l+1}} = 1$, and hence there exists a path $i = i^0, i^2, ..., i^m = j \in f$, contradicting the fact that i and j belong to two different

[5]Note that this walk is not necessarily a path even if the initial walk in h is a path, as the same node j^k can be used several times in the walk.

components in h. Hence all paths between i and j are odd. If there exists an even walk between i and j in g, it must thus involve an odd cycle starting at i or starting at j. Without loss of generality, suppose that there exists an odd cycle starting at i, $i = i^0, ..., i^m = i$, where $m = 2l + 1$ is odd. Consider any even path between i and j, where we index $i = i^m, ...i^r = j$ and $r = 2p$ is even. We construct a path in h between i and j as follows. Because m is odd, we first construct the sequence of connected nodes in h, $i = i^{2l+1}, i^{2l-1}, i^{2l-3}, .., i^1$. Because i is connected in g both to i^1 and i^{m+1}, we then link i^1 to i^{m+1} in h. Now $m + 1$ is even, so we can use the path between i and j to construct a sequence $i^{m+1}, .., i^r = j$ in h. Concatenating the two sequences, we construct a sequence $i, i^{m-2}, ..., i^1, i^{m+1}, .., i^r = j$ in h, contradicting the fact that i and j belong to two different components in h. Hence if h is not connected, there exists a pair of nodes i, j such that all walks between i and j are odd, proving the necessity of the claim.

We now prove the second claim

Claim 6. *For all $i, j \in N$ there exists an even walk between i and j if and only if g is not bipartite.*

Proof. Suppose that g is bipartite with sets A and B. As $N \geq 3$, at least one of the two sets has more than one element. Pick i, j such that $i \in A$ and $j \in B$, then we claim that all walks between i and j must be odd. Any walk between i and j must contain an even number of edges alternating between nodes in A and B and a single edge between a node in A and a node in B. Hence the total number of edges must be odd, proving the necessity of the claim.

Conversely, suppose that there exists a pair of nodes i, j such that all walks between i and j are odd. Consider the sets of nodes $A = \{k|\delta(i, k) \text{ is even}\}$ and $B = \{k|\delta(i, k) \text{ is odd }\}$, where $\delta(i, k)$ denotes the geodesic distance between i and k in the graph. We first claim that if $k \in A$, *all walks between i and k must be even.* Suppose not, then there exist two different walks between i and k, one w_1 which is even (the shortest path between i and k) and one w_2 which is odd. Pick one particular path p between k and j. If this path is odd, then the walk between i and j containing w_2 followed by p is even, contradicting the assumption. If the path is even, then the walk between i and j containing w_1 followed by p is even, contradicting the assumption again. Hence all walks between i and nodes in A are even and all walks between i and nodes in B are odd. Next notice that there cannot be any edge between nodes in A. Suppose by contradiction that there exists an edge between k and l in A, and consider a walk between i and k, w_1 followed by the edge kl. This forms an odd walk between i and l, contradicting the fact that all walks between i and l must be even. Hence, there is no edge between nodes in A and similarly no edge between nodes in B, showing that the graph g is bipartite.

Proof of Proposition 3

Consider a mechanism where all individuals in A are ranked above individuals in B. For any two individuals i and j in A, let $\rho(i) > \rho(j)$ if $i \succ_{\mathbf{T}} j$. If $i \bowtie_{\mathbf{T}} j$ or if the reports on i and j are incompatible, construct an arbitrary ranking by letting $\rho(i) > \rho(j)$ if and only if $i > j$. Similarly, for any two individuals i and j in B, let $\rho(i) > \rho(j)$ if $i \succ_{\mathbf{T}} j$. If $i \bowtie_{\mathbf{T}} j$ or if the reports on i and j are incompatible, let $\rho(i) > \rho(j)$ if and only if $i > j$.

We will show that the mechanism satisfies the three properties. Clearly if i and j are incomparable under two profile types $\mathbf{T}$ and $\mathbf{T}'$, either one belongs to A and the other to B (in which case the mechanism ranks them in the same way under $\mathbf{T}$ and $\mathbf{T}'$), or they belong to the same set, and the mechanism ranks them identically under $\mathbf{T}$ and $\mathbf{T}'$ as it only uses the index to rank them. Hence independence is satisfied.

The mechanism satisfies strategy-proofness, a stronger incentive compatibility notion than ex post incentive compatibility. Consider an individual i in A. Then we claim that if $t^i_{jk} \neq 0$ it must be that both j and k are in B. To see this notice that as g is bipartite it does not contain any triangle. Hence no self-report can be supported by a third individual, and hence $t^i_{ij} = 0$ for all $j \neq i$. The only case where $t^i_{jk} \neq 0$ is thus when $g_{ij}g_{ik} = 1$ and $j, k \in B$. Hence, by changing his report t^i_{jk}, individual i can only affect the ranking of individuals in B. As all individuals in B are ranked below individuals in A, this does not affect the rank of individual i, and hence individual i's ranking is independent of his announcement, proving that the mechanism is strategy-proof.

Finally, notice that by construction, the mechanism ρ achieves an ex post efficient ranking separately on each of the two components A and B. by Proposition 2, the comparison network h is disconnected into two components A and B. Hence the mechanism ρ is also ex post efficient.

Proof of Proposition 4

We consider the same mechanism as in the proof of Theorem 1: We define the mechanism ρ by constructing comparisons. Let r_{ij} denote the comparison between i and j chosen by the planner.

First consider a pair of individuals (i, j) who observe each other, $g_{ij} = 1$. By assumption, there are at least three reports on the ranking of i and j. If all individuals transmit the same report on (i, j), let $r_{ij} = t_{ij}$. If all individuals but one transmit the same report t_{ij} and one individual reports $t'^{ij} = -t_{ij}$, ignore the ranking t'_{ij} and let $r_{ij} = t_{ij}$. In all other cases, let $r_{ij} = 1$ if and only if $i > j$.

Second consider a pair of individuals (i, j) who do not observe each other, $g_{ij} = 0$. If there are at least three individuals who observe i and j, use as above a mechanism such that $r_{ij} = t_{ij}$ if all individuals agree on t_{ij} or only one individual chooses $t'_{ij} = -t_{ij}$, and let $r_{ij} = 1$ if $i > j$ otherwise.

If one or two individuals observe i, j, pick the individual k with the highest index. Consider the vector of announcements $\tilde{T}^{-k}$ where one disregards the announcements of individual k. Let $\succ_{\tilde{T}^{-k}}$ be the binary relation created by letting $t_{ij} = 1$ if and only if $t^l_{ij} = 1$ for all $l \neq k$. If there exists a directed path of length greater or equal to 2 between i and j in $\succ_{\tilde{T}^{-k}}$, and for all directed paths between i and j in $\succ_{\tilde{T}^{-k}}$, $i^0, .., i^L$ we have $t_{i^l i^{l+1}} = 1$, then $r_{ij} = 1$. If on the other hand for all directed paths between i and j in $\succ_{\tilde{T}^{-k}}$, $t_{i^l i^{l+1}} = -1$, then $r_{ij} = -1$. In all other cases, let individual k dictate the comparison between i and j, $r_{ij} = t^k_{ij}$.

Now consider all comparisons r_{ij}. If they induce a transitive binary relation on N, let ρ be the complete order generated by the comparisons. Otherwise, consider all shortest cycles generated by the binary relation r_{ij}. If there exists a single individual i who dictates at least two comparisons in all shortest cycles, individual i is punished by setting $\rho_i = 1$ and $\rho_j > \rho_k$ if and only if $j > k$ for all $j, k \neq i$. If this is not the case, pick the arbitrary ranking where $\rho_i > \rho_j$ if and only if $i > j$.

We now prove that this mechanism satisfies all three conditions. Consider two type profiles $\mathbf{T}$ and $\mathbf{T}'$, and two individuals i and j such that $i \bowtie_{\mathbf{T}} j$ and $i \bowtie_{\mathbf{T}'} j$. Because $\mathbf{T}$ and $\mathbf{T}'$ generate identical truthful reports and result in a transitive partial order, i and j must be ranked at the final completion phase of the mechanism, using the same ranking $\rho(i) > \rho(j)$ if and only if $i > j$. Hence independence holds.

Next consider any pair (i, j) such that $i \succ_{\mathbf{T}} j$. There must exist a sequence of comparisons $(i, i^1, .., i^t, .., i^T, j)$ such that $h_{i^{t-1} i^t} = 1$ and $i^{t-1} \succ_{\mathbf{T}} i^t$. For any of these pairs, we must have $r_{i^{t-1} i^t} = 1$ and hence, because the announcement $\mathbf{T}$ generates a transitive partial order, $\rho(i^{t-1}) > \rho(i^t)$. But this implies that $\rho(i) > \rho(j)$, establishing that the mechanism satisfies ex post efficiency.

Finally, we show that the mechanism is ex post incentive-compatible. Consider individual k's incentive to change his announcement on a link ij when all other individuals tell the truth. If the link ij is supported, this change does not affect the outcome of the mechanism. So consider an unsupported link ij and let individual k be the highest index individual observing i and j. Suppose that all individuals $l \neq k$ announce the truth, so that $\succ_{\tilde{T}^{-k}} = \succ_{T^{-k}}$. We first show that individual k cannot gain by making an announcement which generates cycles in the ranking r_{ij}.

Suppose that the binary relation generated by r_{ij} exhibits a cycle $i^0 i^1 ... i^L$

By the same argument as in the proof of Theorem 1, individual k must dictate at least two comparisons in the cycle. We will show that the initial cycle must contain a cycle of length 3. Suppose that the initial cycle has length greater than or equal to 4. Let ij and lm be two comparisons dictated by individual k. Suppose first that the rank of l is strictly higher than the rank of j. As individual k observes both (i, j) and (l, m), he also observes both i and l. Hence i and l must be compared under r and either $r_{il} = 1$ or $r_{il} = -1$. Now if $r_{il} = -1$, one can construct a shorter cycle by replacing the path $lm..i$ by the path li. If $r_{il} = +1$, one can construct a shorter cycle by replacing the path $ij, , l$ by the path il. Next suppose that $j = l$ so that the two comparisons (i, j) and (l, m) are adjacent in the cycle. Again because individual k observes both i and m, then i and m must be compared under r and either $r_{im} = +1$ or $r_{im} = -1$. If $r_{im} = +1$, one can construct a shorter cycle by replacing ijm with im. If $r_{im} = -1$, one can construct a cycle of length 3 $ijmi$.

We conclude that if the binary relation r exhibits a cycle, there must exist a sub-cycle of length 3, so that all shortest cycles are of length 3. Furthermore, individual k must dictate at least two of the comparisons in all cycles. Hence, individual individual k has no incentive to make an announcement generating a cycle in r, as he will be punished and obtain the lowest rank.

We finally assume that the ranking generated by r is acyclic and show that the rank of individual k must remain the same if he changes his report on any pair (i, j). Notice first that, if k dictates the ranking between i and j, i and j there does not exist a l such that i, j and l can be ranked under T^{-k}. In fact, if i, j and l can be ranked using the reports of individuals in $N \setminus k$, they unanimously rank i and j through a path of length equal or greater than 3 and the mechanism does not let k dictate the choice between i and j. But this implies that whenever k is a dictator over the pair (i_l, j_l) both individuals i_l and j_l are either both ranked above or below k by the reports T^{-k} .

Now let $J = (i_1, j_1), ..., (i_l, j_l), ..(i_L, j_L)$ be the pairs on which k is a dictator and let J^+ denote the set of pairs (i_l, j_l) such that $k \succ_{T^{-k}} i_l, j_l$ and J^- the set of pairs such that $k \prec_{T^{-k}} i_l, j_l$. Let $\mathbf{T'} = (T^{-k}, T'^k)$ the announcement obtained when i changes his report on some of the pairs in J while keeping a transitive partial order. For any m such that $k \succ_{T-k} m$, $k \succ_{\mathbf{T}} m$ and $k \succ_{\mathbf{T'}} m$. Hence $\rho_k(\mathbf{T}) > \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) > \rho_m(\mathbf{T'})$. Similarly, for any m such that $k \prec_{T-k} m$, $k \prec_{\mathbf{T}} m$ and $k \prec_{\mathbf{T'}} m$. Hence $\rho_k(\mathbf{T}) < \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) < \rho_m(\mathbf{T'})$. We also have, for any $m \in J^+$, $\rho_k(\mathbf{T}) > \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) > \rho_m(\mathbf{T'})$. For any $m \in J^-$, $\rho_k(\mathbf{T}) < \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) < \rho_m(\mathbf{T'})$. Next consider m such that $k \bowtie_{T^{-k}} m$ and $m \notin J$. If $m \prec_{T^{-k}} i_l$ for some $i_l \in J^+$, then $k \succ_{\mathbf{T}} m$ and $k \succ_{\mathbf{T'}} m$ so that $\rho_k(\mathbf{T}) > \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) > \rho_m(\mathbf{T'})$. Similarly, if $m \succ_{T^{-k}} i_l$ for some $i_l \in J^-$, then $k \prec_{\mathbf{T}} m$ and $k \prec_{\mathbf{T'}} m$ so that $\rho_k(\mathbf{T}) < \rho_m(\mathbf{T})$ and $\rho_k(\mathbf{T'}) < \rho_m(\mathbf{T'})$. Finally, if $m \succ_{T^{-i}} i_l \forall i_l \in J^+, m \prec_{T^{-k}} i_l \forall i_l \in J^-$ and $k \bowtie_{T^{-k}} m$, then $k \bowtie_{\mathbf{T}} m$ and $k \bowtie_{\mathbf{T'}} m$. Whenever $k \bowtie_{\mathbf{T}} m$ and $k \bowtie_{\mathbf{T'}} m$, then the ranking between k and m is independent of the type profile. Hence, in all cases the ranking between k and m is identical under $\mathbf{T}$ and $\mathbf{T'}$. This argument completes the proof that the mechanism satisfies ex post incentive compatibility.

Proof of Proposition 5

We first establish the following simple general claim:

Claim 7. *If ρ is strategy-proof, $\rho_i(T^i, \hat{T}^{-i}) = \rho^i(T'^i, \hat{T}^{-i})$ for all $i, T^i, T'^i, \hat{T}^{-i}$.*

Proof. Suppose by contradiction that there exists $i, T^i, T'^i, \hat{T}^{-i}$ such that $\rho_i(T^i, \hat{T}^{-i}) > \mu^i(T'^i, \hat{T}^{-i})$. Let T'^i be the true type of individual i. Then, individual i has an incentive to announce T^i, contradicting the fact that ρ is strategy-proof.

Consider next two vectors of types:

- $\mathbf{T}_1$: $t_{ij} = t_{jk} = t_{ik} = 1$
- $\mathbf{T}_2$: $t_{ij} = -1, t_{jk} = 1, t_{ik} = -1$

As the mechanism is ex post efficient, it must assign ranks $\rho_i(\mathbf{T}_1) = 3, \rho_j(\mathbf{T}_1) = 2, \rho_k(\mathbf{T}_1) = 1, \rho_i(\mathbf{T}_2) = 1, \rho_j(\mathbf{T}_2) = 3, \rho_k(\mathbf{T}_2) = 2$.

Now let t^i denote the announcement $t^i_{ij} = t^i_{jk} = t^i_{ik} = 1$ and t'^i the announcement $t^i_{ij} = -1, t^i_{jk} = 1, t^i_{ik} = -1$. By Claim 7,

$$\begin{aligned}
\rho_i(t'^i, t^j, t^k) &= \rho_i(t^i, t^j, t^k) &&= 3,\\
\rho_j(t^i, t'^j, t^k) &= \rho_j(t^i, t^j, t^k) &&= 2.\\
\rho_k(t'^i, t'^j, t^k) &= \rho_k(t'^i, t'^j, t'^k) &&= 2.
\end{aligned}$$

Hence we conclude that, at (t'^i, t'^j, t^k) either $\rho_i = 3, \rho_j = 1$ or $\rho_i = 1, \rho_j = 3$. But $\rho_j = 3$ is impossible, as, by claim 7, $\rho_j(t'^i, t'^j, t^k) = \rho_j(t'^i, t^j, t^k)$ and $\rho_j(t'^i, t^j, t^k) \neq \rho_i(t'^i, t^j, t^k) = 3$. Hence we conclude that

$$\rho_i(t'^i, t'^j, t^k) = 3, \rho_j(t'^i, t'^j, t^k) = 1, \rho_k(t'^i, t'^j, t^k) = 2. \tag{1}$$

A similar reasoning shows that

$$\rho_j(t'^i, t^j, t'^k) = 3,$$

and hence either $\rho_i = 2, \rho_k = 1$ or $\rho_i = 1, \rho_k = 3$ at (t'^i, t^j, t'^k). But $\rho_i(t'^i, t^j, t'^k) = \rho_i(t^i, t^j, t'^k) \neq \rho_k(t^i, t^j, t'^k) = \rho_k(t^i, t^j, t^k) = 1$. So we conclude that

$$\rho_i(t'^i, t^j, t'^k) = 2, \rho_j(t'^i, t^j, t'^k) = 3, \rho_k(t'^i, t^j, t'^k) = 1. \tag{2}$$

Now, $\rho_j(t'^i, t^j, t^k) = \rho_j(t'^i, t'^j, t^k)$. By equation 1,

$$\rho_j(t'^i, t'^j, t^k) = 1$$

,

so that $\rho_j(t'^i, t^j, t^k) = 1$. As $\rho_i(t'^i, t^j, t^k) = \rho_i(t^i, t^j, t^k) = 1$,

$$\rho_k(t'^i, t^j, t^k) = 2.$$

Similarly, $\rho_k(t'^i, t^j, t^k) = \rho_k(t'^i, t^j, t'^k)$ and by equation 2,

$$\rho_k(t'^i, t^j, t'^k) = 1$$

so that

$$\rho_k(t'^i, t^j, t^k) = 1.$$

establishing a contradiction.

Proof of Proposition 6

If individuals strictly prefer being ranked at $\rho(i) = 2$ to being ranked at $\rho(i)$, the necessity part of the proof of Theorem 1 shows that whenever there exists a pair of individuals who are not observed by a third individual, there cannot exist an ex post incentive-compatible and efficient mechanism.

Conversely, if individuals are indifferent between being ranked at $\rho(i) = 1$ and $\rho(i) = 2$, let $\rho(i) = 1$ and $\rho(j) = 2$ whenever i and j are the only two individuals observing the ranking between i and j and $t^i_{ij} \neq t^j_{ij}$. This guarantees that individuals have no incentive to send conflicting reports, and hence that this mechanism, completed by the mechanism constructed in the sufficiency part of Theorem 1, satisfies ex post incentive compatibility and efficiency.

Proof of Proposition 7

Consider a vector of announcements where all three individuals agree on $t_{13} = -1, t_{23} = -1, t_{12} = 1$. By ex post efficiency, $\rho(1) = 2, \rho(2) = 1, \rho(3) = 3$. We claim that ex post group-incentive compatibility implies that, whenever individual 3 announces $t^3_{13} = -1, t^3_{23} = -1, t^3_{12} = 1$, the rank of individual 1 must be different from 3. If that were not the case, there would exist an announcement (t'^1, t'^2) for individuals 1 and 2 resulting in a rank $\rho(1) = 3 > 2, \rho(2) \geq 1$, contradicting ex post group-incentive compatibility. By a similar reasoning, whenever individual 1 announces $t^1_{12} = 1, t^1_{13} = 1, t^1_{23} = 1$, the rank of individual 2 must be different from 3. Finally, when individual 2 announces $t^2_{12} = -1, t^2_{23} = 1, t^2_{13} = -1$, the rank of individual 3 must be different from 3.

So consider the announcement $t^1 = (t^1_{12} = 1, t^1_{13} = 1, t^1_{23} = 1), t^2 = (t^2_{12} = -1, t^2_{23} = 1, t^2_{13} = -1), t^3 = (t^3_{13} = -1, t^3_{23} = -1, t^3_{12} = 1)$. For this announcement, neither of the three individuals can be in position 3, a contradiction which completes the proof of the Proposition.