



# Inducing Peer Pressure to Promote Cooperation

Ankur Mani<sup>1</sup>, Iyad Rahwan<sup>2,3</sup> & Alex Pentland<sup>1</sup>

SUBJECT AREAS:

CLIMATE-CHANGE  
POLICY

SOCIOECONOMIC SCENARIOS

PSYCHOLOGY AND BEHAVIOUR

APPLIED PHYSICS

<sup>1</sup>Massachusetts Institute of Technology, Cambridge MA, USA, <sup>2</sup>Masdar Institute of Science and Technology, Abu Dhabi, UAE, <sup>3</sup>University of Edinburgh, Edinburgh, UK.

Received  
21 January 2013

Accepted  
9 April 2013

Published  
26 April 2013

Cooperation in a large society of self-interested individuals is notoriously difficult to achieve when the externality of one individual's action is spread thin and wide on the whole society. This leads to the 'tragedy of the commons' in which rational action will ultimately make everyone worse-off. Traditional policies to promote cooperation involve Pigouvian taxation or subsidies that make individuals internalize the externality they incur. We introduce a new approach to achieving global cooperation by localizing externalities to one's peers in a social network, thus leveraging the power of peer-pressure to regulate behavior. The mechanism relies on a joint model of externalities and peer-pressure. Surprisingly, this mechanism can require a lower budget to operate than the Pigouvian mechanism, even when accounting for the social cost of peer pressure. Even when the available budget is very low, the social mechanisms achieve greater improvement in the outcome.

Correspondence and requests for materials should be addressed to A.P. (sandy@media.mit.edu)

Cooperation in large societies of self-interested individuals is a crucial, yet extremely difficult goal to achieve<sup>1</sup>. Some of the most important problems in modern society, such as pollution, global warming, rising health care costs and insurance, arise from the inability to achieve cooperation over a large scale. If everyone does their small part in reducing pollution, consuming responsibly, maintaining a healthy lifestyle and contribute to the health insurance, these problems will be solved. The *tragedy of the commons* occurs when multiple individuals, acting rationally in their own self-interest, will ultimately deplete a common resource, to the detriment of everybody<sup>2</sup>. The cause of the tragedy is that the *negative externality* from any individual's non-cooperative action is experienced by the large society, yet the benefit is borne entirely by the individual. For example, when someone buys an SUV with high CO<sub>2</sub> emission as opposed to a hybrid vehicle with low CO<sub>2</sub> emission, the full cost of pollution (e.g. in terms of health care cost and environmental effects) is borne by the whole society. The cost incurred by the individual is small compared to the benefit, and so he *free rides*.

One traditional solution to the problem of cooperation in large societies is enforcing *quotas* that limit the production of negative externalities (e.g. cap on carbon emission). An alternative, market-based approach is *Pigouvian taxation*<sup>3</sup>. The idea is to bring up the cost of one's actions to a level that accounts for the externality incurred. For example, a carbon tax may be added to the price of fuel in order to account for the true cost of consumption on the rest of society (and used to plant trees, say). By taking this true cost into account, the individual no longer has incentive to free-ride, and will only purchase the SUV if it is worth the entire cost to the individual. Governments also often use *Pigouvian subsidies* (e.g. subsidize the price of solar panels) to encourage activities with positive externalities. In effect, these policies tax everyone in the society and redistribute it as subsidies to enforce cooperation.

The outcomes of such policies may be socially sub-optimal for two reasons. Firstly, the Coasian argument<sup>4</sup> fails due to the presence of significant transaction cost<sup>5</sup> and so a simple redistribution does not achieve a Pareto-efficient outcome. Secondly, these policies assume that the society consists of a population of independent individuals and discounts the fact that individual decisions are influenced by the interactions with the peers in the society.

There is growing evidence of the power of social influence in general, and peer pressure in particular, in promoting cooperative behavior. In evolutionary biology, mechanisms like direct reciprocity, network reciprocity, and group selection all work by exploiting locality of interaction<sup>6</sup>. When people interact in small groups in experimental public goods games (PGG), the ability to punish or reward peers (even at a cost) has been shown to promote cooperative behavior<sup>7,8</sup>. Similarly, economic models of peer pressure show that pressure acts as a mechanism for improving the social welfare of the group<sup>9</sup> and is conducive of successful business partnerships<sup>10</sup>. In Micro Finance Institutions, stronger social ties can increase the likelihood of repayment of joint-liability loans, by facilitating monitoring and enforcement<sup>11</sup>. The effectiveness of peers on loan repayment has been demon-



strated even in the absence of joint liability<sup>12</sup>. In relation to natural resource usage, strong community identification is instrumental to preventing overuse of water resources<sup>13</sup>, and so do other community-based incentives<sup>14</sup>. Top-down rules have been blamed for the degraded inshore ground fishery in Maine, in contrast to the Maine lobster fishery, which has been governed by informal community-based institutions and yielded much higher levels of compliance with sustainable resource use<sup>1</sup>.

What the above examples highlight is that many mechanisms for promoting cooperative behavior work *locally*. An individual free-riding on immediate social ties (e.g. causing bad smell in the neighborhood by burning paper for cooking) has much to suffer: complaint, loss of reputation, social exclusion, etc. In contrast, in the 'tragedy of the commons' scenario (e.g. CO<sub>2</sub> emission from an SUV), most of the externality is felt by individuals who are not peers and have no way to exert peer-pressure. Yet the cost of exerting peer-pressure is higher for any given peer than the externality experienced. Consequently, the peers do not exert sufficient pressure to nudge the individual to cooperate. Indeed, centralized sanctioning institutions exist precisely to promote cooperation in larger human groups, in which peer punishment is not sufficient<sup>15</sup>.

Against this background, we propose a new set of *social mechanisms* for policy makers to address the problem of externalities in a networked society, in which externalities are global, but interactions are local. The mechanisms work by inducing peer-pressure to promote cooperation.

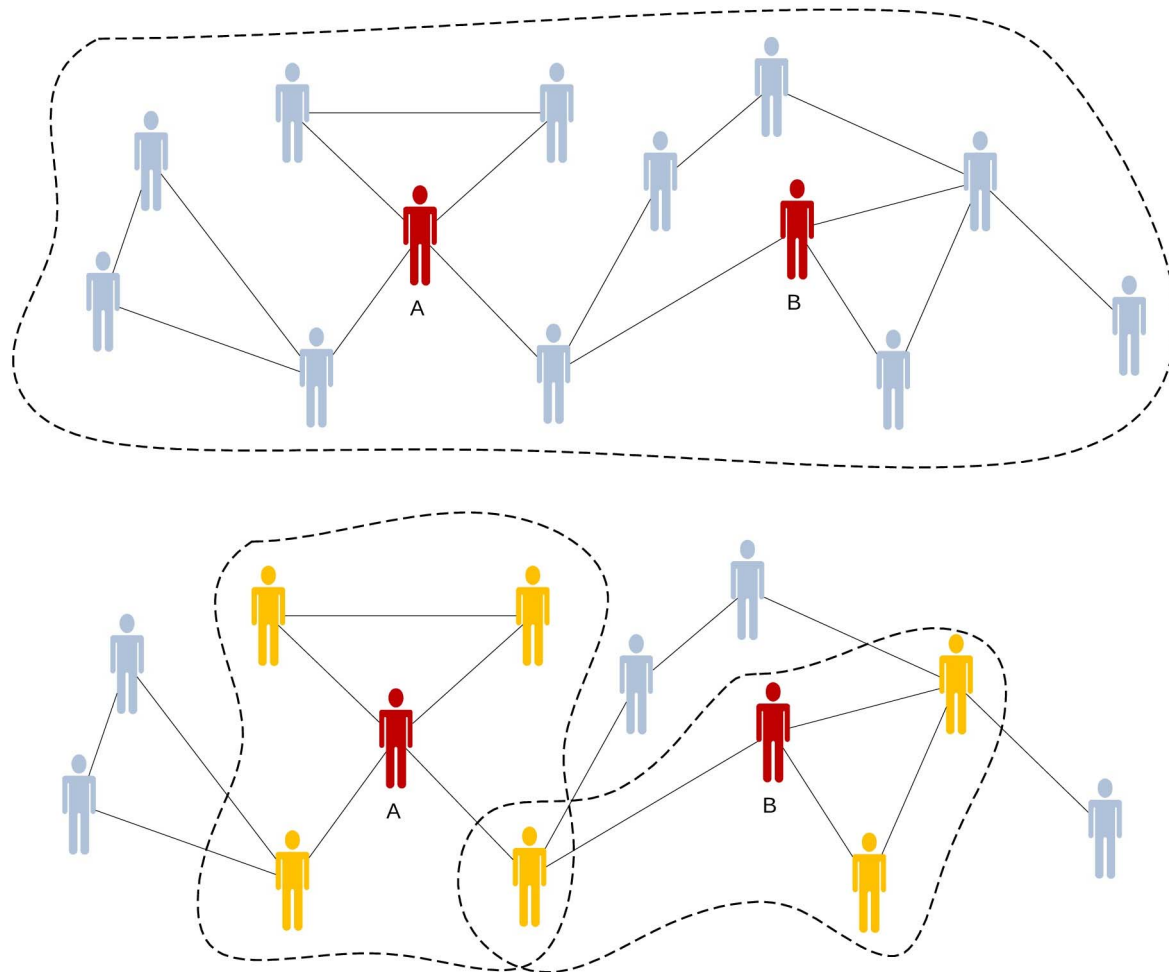
## Results

**Overview of mechanism.** The main idea behind the mechanism is illustrated in Figure 1. In contrast with the Pigouvian approach, which focuses on the individual causing the externality, our mechanism focuses on their peers in the social network. The idea is to incentivize agent A's peers to exert (positive or negative) pressure on A. In reality, such peer pressure may take many forms. In the energy consumption scenario, examples of positive pressure include giving useful energy-saving advice to a neighbor, or giving the neighbor a hand at installing a solar panel. Examples of negative pressure include inducing guilt or shame<sup>10</sup>, or slashing the tires of the neighbor's new SUV<sup>7</sup>.

Regardless of its form, peer pressure requires costly effort on behalf of agent A's peers. A peer often has little incentive to exert such pressure, since the resulting effect, in terms of reduction in negative externality experienced by the given peer, is very low. Our mechanism can be summarized by the following question: *If we reward the peers of agent A, can we encourage them to exert more pressure on A to reduce the negative externality? And is this policy efficient compared to Pigouvian policies?*

Our main insight is that by targeting the individual's peers, peer pressure can *amplify* the desired effect on the target individual. That is, under certain conditions, the resulting reduction in negative externality can be larger, given an identical subsidy budget.

Our results hold for positive and negative externalities, positive and negative peer pressure, as well as subsidy and taxation. For



**Figure 1** | Top: Individuals A and B (red) actions cause a global externality on the rest of society (blue). This behavior can be encouraged/discouraged by internalizing the externality in the form of subsidy or tax on A and B directly; Bottom: Localizing externalities to the peers (yellow) of individuals A and B incentivizes their respective peers to use peer influence to encourage/discourage the behavior causing the positive/negative externality, respectively.



simplicity, in the remainder of the paper, we focus on a single scenario. An individual chooses a level of action (e.g. consumption of electricity) which causes negative externalities (e.g. pollution). An external entity (e.g. municipal government) attempts to encourage the individual to reduce their action level through a direct reward scheme (e.g. subsidized solar panels). At the same time, the individual's peers are able to exert negative pressure on the individual.

We study a joint strategic model of externalities and peer pressure in social networks where agents take actions that exert externalities on the whole network and may apply costly peer pressure on their peers. We model this as a two stage game in which, in the first stage, the individuals in the network choose the amount of peer pressure to put on each of their peers. In the second stage, the individuals choose their action that puts externality on the whole society. We study the sub-game perfect equilibrium of this game, show existence, conditions for uniqueness of actions across different equilibria, and how the peer-pressure is distributed in the social network in the different equilibria. We find that, in the equilibrium of this game, only the peers who feel the highest externality apply pressure. Furthermore, the pressure felt by any individual in the network is the same in all the equilibria. This pressure yields some improvement in social surplus, but may not be optimal.

With these characterization results, we explore how, using information about the structure of the social network, optimal social surplus can be achieved using carefully designed social mechanisms. Furthermore, we show that the social mechanism achieves the optimal outcome at a lower budget and total cost than the Pigouvian mechanism.

Social mechanisms are superior for two reasons: (i) When all the externalities are internalized as in the Pigouvian mechanism then there is no peer pressure on the agent creating the externality, and thus requires additional subsidies; and (ii) When the marginal cost of exerting peer pressure is lower than the marginal externality on the whole society times the marginal response to peer-pressure then the effect of subsidies is amplified in the social mechanism. This amplification increases with the strength of the relationships between the peers, and is inversely proportional to the cost of exerting peer pressure.

We anticipate two applications of such mechanisms: (i) public policy for reducing global externalities such as pollution, and (ii) revenue maximization for products with network externalities such as collaborative search engines, or social recommendations.

**Model of externalities.** We now present the formal model. Consider a set of agents  $N$  in a social network  $S = (N, E)$ , where  $E \subseteq N \times N$ . Let  $Nbr(i) = \{j : (i, j) \in E\}$  be the set of peers of the actor  $i$ . We assume that the social network is sparse and agents have at most  $K$  peers. An agent  $i \in N$  takes an action  $x_i \in \mathbb{R}_+$  (e.g. corresponding to units of electricity consumed), and define  $\mathbf{x} \in \mathbb{R}_+^{|N|}$  be an action profile of all agents. Each agent  $i$  experiences raw utility from its action defined by the function  $u_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ . We assume  $u_i$  is a twice differentiable and strictly concave raw utility function with a unique maximum and lower bounded first derivative that approaches infinity as the action approaches zero. This is a very natural case (see SI Appendix for more).

In the standard model of externalities [16, Ch. 11], it is assumed that the utility of actor  $i$  depends both on the raw utility of its own action as well as the externalities experienced due to the actions of others. The latter can be captured by a function  $v_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ , which is strictly convex and increasing, and captures the externality experienced by  $i$  due to the aggregate action of other actors in the population. Therefore, the total utility of actor  $i$ , given its own action  $x_i$  and the action,  $\mathbf{x}_{-i}$ , of other agents,  $N \setminus \{i\}$ , is defined as  $U_i(x_i, \mathbf{x}_{-i}) := u_i(x_i) - v_i(\sum_{j \neq i} x_j)$ . Social surplus is defined as the sum of utilities achieved by all agents:  $S(\mathbf{x}) := \sum_{i \in N} U_i(x_i, \mathbf{x}_{-i})$ . We choose a linear social surplus function so that the social surplus maximization selects

a unique Pareto-efficient outcome, among the infinitely many Pareto-efficient outcomes, that can be compared under different mechanisms.

We will denote,  $\mathbf{x}^*$ , as the action profile at equilibrium, and  $\mathbf{x}^\circ$ , the action profile that maximizes social surplus. Note that  $\mathbf{x}^\circ$  is also a Pareto-efficient action profile. At optimal action  $\mathbf{x}^\circ$ , for all agents  $i$ , the raw marginal utility is equal to the marginal externality on the

whole society, i.e.-  $u'_i(x_i^\circ) = \sum_{j \neq i} v'_j(\sum_{k \neq j} x_k^\circ)$ . In the standard

model of externalities, we know that at the equilibrium, the agents take action that is higher than the socially optimal action, that is  $\mathbf{x}^* > \mathbf{x}^\circ$ <sup>16</sup> and therefore the social surplus at the equilibrium is sub-optimal. This is the essence of the Tragedy of the Commons.

**Externalities with peer pressure.** Consider, now, that actors have the ability to exert peer pressure on their peers in the social network. We denote the peer-pressure profile by the matrix  $\mathbf{p} \in \mathbb{R}_+^{N \times N}$ , where the element  $p_{ij}$  is the peer-pressure exerted by the agent  $i$  on her peer  $j$ . Note that if  $i$  and  $j$  are not peers in the social network, then  $p_{ij} = 0$ . A similar model was proposed by Calvo-Armengol and Jackson<sup>9</sup>. In their model, there is no network constraint and all agents can exert peer-pressure on all other agents. Also the model deals with binary action space as opposed to the infinite action space in our model.

The utility of an actor then takes the following extended form:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{p}) = u_i(x_i) - v_i\left(\sum_{j \neq i} x_j\right) - \left(\sum_{j \in Nbr(i)} p_{ji}\right)(x_i - x_i^\circ) - \left(\sum_{j \in Nbr(i)} p_{ij}\right)c$$

Thus, in addition to the raw utility of action and the externality, actor  $i$  also experiences potential disutility that is bilinear in the total pressure from the peers and  $i$ 's own action. An individual's action and the peer pressure on the individual enter as strategic substitutes in the individual's utility. The higher  $i$ 's action is, the more salient the effect of the pressure becomes. Agent  $i$  also incurs a cost  $c \cdot p_{ij}$  should it wish to exert pressure on neighbor  $j$ , where  $c$  is the marginal cost of exerting such pressure, and  $p_{ij}$  is the amount of pressure. Although we consider that the marginal cost of exerting peer pressure is identical for all pairs of peers, the qualitative nature of our results do not change with different costs. We study the externalities model with peer-pressure as a two-stage game. In the first stage, actors choose the amount of peer pressure they wish to exert on their peers. In the second stage, actors observe pressure on themselves and then choose their action as a response to the observed pressure. Given pressure  $\mathbf{p}$ , we denote  $x_i^*(\mathbf{p})$  as the optimal response to the pressure for agent  $i$ . The optimal response is unique and the marginal raw utility for actor  $i$  at the optimal response is equal to the total pressure exerted on  $i$ , i.e.  $u'_i(x_i^*(\mathbf{p})) = \sum_{j \in Nbr(i)} p_{ji}$ . As the pressure on an agent increases, the optimal response decreases and the optimal response is convex in the total pressure on the agent. The marginal response by any agent  $i$  to pressure  $p_{ji}$  from any of her peers  $j$  is the reciprocal of the curvature of the raw utility function at the optimal response, i.e.-  $\frac{\partial x_i^*(\mathbf{p})}{\partial p_{ji}} = \frac{1}{u''(x_i^*(\mathbf{p}))}$ .

**Assumptions.** As in the Tragedy of the Commons, the scenario we are interested in is when the the externality caused rises much slower than the raw utility function and any one agent's marginal change in action has very small effect on the marginal externality experienced by any other agent. In this case, the large externality is due to higher aggregate action of all agents. We are interested in the scenario, where the marginal cost of exerting pressure is neither too high that no one cares to exert any pressure, nor too low that pressure is so high that everyone takes the socially optimal action. Therefore, we make the



following assumptions about the marginal cost of exerting peer-pressure.

Firstly, we are interested in situations in which the cost of peer-pressure is not too low. Formally, the marginal cost of exerting peer-pressure is higher than the ratio of the marginal externality of any agent,  $j$ , to the curvature of the raw utility of any of her peers,  $i$ , when

the action profile is socially optimal  $\mathbf{x}^o$ , i.e.-  $c > \left| \frac{v'_j \left( \sum_{k \neq j} x_k^o(\mathbf{p}) \right)}{u''_i(x_i^o(\mathbf{p}))} \right|$  for all

peers  $i, j$ . In the absence of this condition, optimal pressure is already present, everyone in society is well-behaved (social surplus is optimal), thus eliminating any need for intervention.

Secondly, we are interested in situations in which, at least for some agents, peer-pressure is worth its cost –if the cost of exerting peer-pressure is too high, then no one will exert peer-pressure on anyone. Formally, for at least one agent  $i$  and one of her peers  $j$ , the marginal cost of exerting pressure for  $j$  on  $i$  is lower than the ratio of the marginal externality experienced by  $j$  to the curvature of the raw

utility of  $i$  when the action profile is  $\mathbf{x}^*$ , i.e.-  $c < \left| \frac{v'_j \left( \sum_{k \neq j} x_k^*(\mathbf{p}) \right)}{u''_i(x_i^*(\mathbf{p}))} \right|$ .

This means that  $j$ 's pressure can yield some benefit to her by influencing  $i$ 's behavior.

The above conditions on the cost of peer pressure are satisfied in scenarios such as the Tragedy of the Commons, in which peer pressure can potentially act as a conduit for social welfare improving behavior, but it is too costly to apply sufficient pressure. For example, suppose agent  $i$  is a smoker, paying little regard to others around him, while contributing to the broader problem of lung cancer. From the perspective of his neighbor  $j$ , exerting peer pressure by expressing disapproval or asking him to smoke outside can yield a real benefit in terms of reduced discomfort to  $j$ . Yet, it is sufficiently costly (read socially awkward) for  $j$  to exert enough pressure to get  $i$  to put out the cigarette, let alone quit. This causes  $j$  to refrain from pushing further. This is also the case with problems of externalities such as pollution. The change in externality felt by any one agent (who's putting pressure) due to the reduction in one other agent's pollution level (e.g. replacing his SUV with a hybrid vehicle) is much smaller than the gain in utility for the agent creating the pollution (and enjoying his SUV). Yet, if a large number of agents reduce their pollution level simultaneously, then the reduction in the externality felt by any one agent (reduced pollution) may be much larger than the loss in the utility of any one agent changing her pollution level (replacing the SUV).

Finally, we assume that the society is large enough that the marginal cost of exerting peer-pressure is less than  $\frac{1}{2K}$  times the ratio of marginal externality felt by the whole society and the curvature of the raw utility function for any agent  $i$ , at the socially optimal action profile  $\mathbf{x}^o$ , i.e.-  $c < \frac{1}{2K} \left| \frac{\sum_{j \neq i} v'_j \left( \sum_{k \neq j} x_k^o(\mathbf{p}) \right)}{u''_i(x_i^o(\mathbf{p}))} \right| = \frac{1}{2K} \left| \frac{u'_i(x_i^o(\mathbf{p}))}{u''_i(x_i^o(\mathbf{p}))} \right|$ .

This relates the cost of peer-pressure to the relative locality in interaction. This means that the global externality is sufficiently larger than the local cost (see SI Appendix for the formal details and proofs of all results).

This last assumption highlights that the social network structure plays a role in our results. Social mechanisms are suitable for large networked societies that are sparse, meaning that an individual has very few peers as compared to the total number of people in the society. Most large modern societies demonstrate such structure.

**Equilibrium with peer pressure.** We now study the sub game perfect equilibria in the externalities model with peer-pressure. Although there may be multiple possible subgame perfect equilibria, the total peer-pressure on any agent  $i$  and her action is same in all the subgame perfect equilibria under the following condition: for any pair of peers,

$i, j$ , the elasticity of the marginal externality felt by  $j$  is lower than the elasticity of the curvature of the raw utility of  $i$  with respect to the

agent  $i$ 's action, i.e.-  $\frac{\partial \log v'_j \left( \sum_{k \neq j} x_k \right)}{\partial x_i} < \left| \frac{\partial \log |u''_i(x_i)|}{\partial x_i} \right|$ , for all

action profiles that are not strictly dominated (see SI Appendix for details). This condition suggests that the marginal externalities grow slower than the decay of marginal curvature of the raw utility function. We will assume that this condition is satisfied for the rest of the paper. In the sub-game perfect equilibrium of this game<sup>17</sup>, we observe that the equilibrium action (action) profile in the game with peer-pressure is lower than the action-profile in the game without peer-pressure. This highlights the potential benefit of peer pressure.

**Theorem 1.** The action profile in any subgame perfect equilibrium in the two-stage game with peer-pressure is strictly lower than the equilibrium action profile in the game without peer-pressure.

Under assumption 1, there is positive peer-pressure on all agents in any sub game perfect equilibria and hence the optimal response action for each agent is lower than the optimal response without peer-pressure. We also observe from the application of KKT conditions that in the equilibria, not all agents exert pressure on all their peers (see SI Appendix for details). Assume  $\mathbf{p}^*$  is the peer-pressure profile in an equilibrium.

(i) Each agent feels pressure only from the peers who have the highest marginal externality at the equilibrium action, i.e.-  $p_{ji} > 0$  for the peers  $i, j$  only if  $v'_j \left( \sum_{k \neq j} x_k^*(\mathbf{p}^*) \right) \geq v'_l \left( \sum_{k \neq l} x_k^*(\mathbf{p}^*) \right)$  for all other peers  $l \in Nbr(i)$ .

(ii) Each agent puts pressure only on her peers who have the smallest curvature of the raw utility function at the equilibrium action, i.e.-  $p_{ji} > 0$  for  $i \in Nbr(j)$  only if  $|u''_i(x_i^*(\mathbf{p}^*))| \leq |u''_k(x_k^*(\mathbf{p}^*))|$  for all  $k \in Nbr(j)$ .

While the presence of peer-pressure causes lower individual action, the peer-pressure in equilibrium is not sufficient to bring the action down to the optimal level. Next, we investigate how various interventions affect this.

**Pigouvian mechanism (direct reward).** In the standard Pigouvian mechanism, agents are rewarded for reduced action. The reward given to any agent  $i \in N$  for her action  $x_i$  is  $r_i(x_i) = u'_i(x_i^o)(x_i^* - x_i)$ . The utility function of agent  $i$  under the Pigouvian mechanism therefore has the reward as an extra term. While this reward incentivizes reduction in individual action, it removes any incentive to exert peer pressure so there is no peer-pressure in the equilibrium.

**Proposition 1.** There is no peer-pressure on any agent in equilibrium under Pigouvian mechanism.

This is because under Pigouvian mechanism when the marginal reward for any agent is equal to the externality produced by the agent at the socially optimal action profile, the agents completely internalize the externality they produce and in the equilibrium each agent takes the socially optimal action. Following assumption 2, then the marginal cost of exerting peer-pressure is more than the marginal benefit for any agent as a result of the marginal reduction in externality and therefore no agent exerts peer-pressure on her peers (see SI Appendix for details). So the Pigouvian mechanism does not use any social capital.

**Social mechanisms (rewarding the peers).** As described earlier, the social mechanism rewards individuals for their peers' low action, in effect subsidizing the cost of peer pressure they incur. There are many possible reward structures for creating social mechanism. We discuss here one structure in which the reward is given to agent  $i$  as a result of her peer, agent  $j$ 's action  $x_j$ . Formally,





$r_{ji} : \mathbb{R}_+ \rightarrow \mathbb{R}$  is strictly decreasing in  $x_j$  (the more  $j$  consumes, the less reward  $i$  gets), and  $r_{ji}(x_j)$  is the reward given to the actor  $i$  for the action  $x_j$  by a neighboring actor  $j$ . Since  $i$  can have multiple peers, his utility incorporates aggregate social reward  $\sum_{j \in N, j \neq i} r_{ji}(x_j)$ .

What might be a suitable approach to allocating such social reward? We would like to identify a reward function which has the following properties:

1. The reward must be simple. We consider reward functions with constant marginal reward (i.e. affine reward functions).
2. A subgame perfect equilibrium of the two-stage game should exist.
3. Equilibrium action should be optimal.
4. Each peer gets rewarded for an agent's reduced action.
5. Budget for rewards should be minimized over the set of reward functions that satisfy the above conditions.

It turns out that a fairly simple reward function satisfies these conditions. Changing the conditions 1–5 may give different reward functions. Note that this reward has a component that depends upon the consumer and a component that depends upon the neighbor.

**Theorem 2.** The following social reward function, to agent  $i$  based on the action of agent  $j$ , satisfies conditions (1–5):

$$r_{ji}(x_j) = (\alpha_j + \beta_i)(x_j - x_j^*)$$

where  $\alpha_j = cu''_j(x_j^0)$  and  $\beta_i = v'_i(\sum_{k \neq i} x_k^0)$

The marginal reward is such that each agent has an incentive to increase pressure on peers that take a higher action than socially optimal level and reduce pressure on peers that take lower action than social optimal level.

**Efficiency and cost of social reward.** Our most striking result is that (under assumption 3) the social mechanism is significantly more efficient than the Pigouvian mechanism. To establish this, we first compare the budgets needed by the two mechanisms in order to reach optimal action.

**Theorem 3.** The budget for the rewards in the Pigouvian Mechanism is at least twice the budget for the rewards in the social mechanism.

While this result is very encouraging, a more comprehensive comparison must account for the total cost of the two mechanisms. Two additional costs needs to be accounted for. First, there is an aggregate *social cost* of exerting pressure on peers, incurred by all those agents who choose to do so. This can be thought of as a *loss of social capital*. While a single definition of the term is still not established<sup>18</sup>, applying peer pressure can be seen as utilizing some existing *social capital* that one possesses<sup>19</sup>.

Secondly, there is a *redistribution loss* incurred by both mechanisms, to account for the cost of distributing rewards (or of taxation). We assume that this is simply a fraction  $\delta$  of the reward given.

It turns out that the social cost incurred due to peer-pressure in the equilibrium under social mechanism is smaller than the redistribution loss in the equilibrium under Pigouvian mechanism when the redistribution factor is sufficiently high.

**Theorem 4.** The loss in social capital in the equilibrium under the social mechanism is lower than the redistribution loss in the equilibrium under the Pigouvian mechanism for all redistribution factors

$$\delta \geq \frac{1}{2K}$$

Based on this result, we can characterize conditions under which the cost of the social mechanism is lower than the Pigouvian mechanism.

**Corollary 1.** The total loss in the equilibrium under the social mechanism including the loss in the social capital and the redistribution loss is lower than the the redistribution loss in the equilibrium under the Pigouvian mechanism for distribution factors  $\delta \geq \frac{1}{K}$ .

This result shows that the social mechanism is superior, even when accounting for the cost to society, and not just considering the cost to the party operating the mechanism. Often, the budget available is lower than the required budget for rewards. Under these circumstances, the marginal reward is kept low and the whole externality created by the agents may not be completely internalized under the Pigouvian mechanism or completely localized under the social mechanism. A corollary is that when the marginal reward rate is very low, the action profile in equilibrium under the social mechanism is lower than the action profile in equilibrium under the Pigouvian mechanism. This is important for two reasons. Firstly, since the raw utility functions and the externality functions are often difficult to estimate due to limited availability of accurate data, and since the budget is often insufficient, for most practical purposes, the marginal reward rate is low. Secondly, it can be tested experimentally. Our results provide three testable predictions, namely:

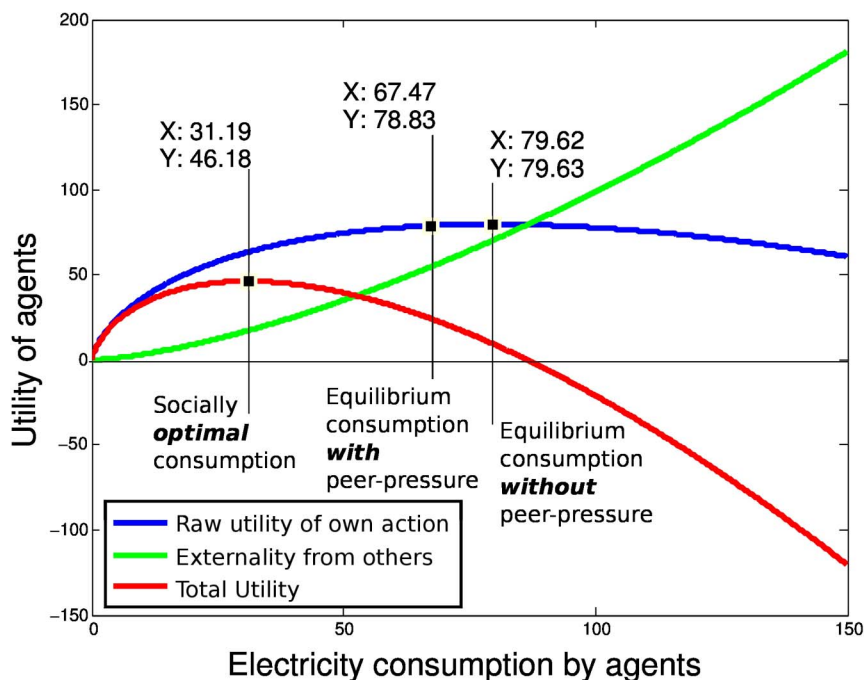
1. Following theorem 1, the equilibrium action profile in the game with peer-pressure has a higher social surplus than the equilibrium action profile in the game without peer-pressure.
2. As a corollary of theorem 4, for low marginal rewards, the equilibrium action profile under the social mechanism has a higher social surplus than the equilibrium action profile under the Pigouvian mechanism.
3. When the marginal cost of exerting peer-pressure decreases (i.e.-peers are closer friends), the social surplus in the equilibrium under social mechanism increases.

**Illustrative example: promoting energy efficiency.** Consider a homogeneous set of agents  $N = \{1, \dots, 100\}$ , each agent has 10 peers in the social network. Each agent consumes electricity priced at 4 per unit. The raw utility function of all agents is  $u_i(x_i) = 12x_i^{0.8} - 4x_i$  for all  $i \in N$ , and the externality function of all agents is  $v_i(y) = 0.0001(y)^{1.5}$  for all  $i \in N$ . The marginal cost of exerting pressure is  $c = 1$  per unit. Assume that the redistribution loss is  $\delta = 0.1$  per unit of reward. Figure 2 shows how the raw utility of consumption, externality and total utility of each agent changes with the increase in consumption, assuming each agent has the same consumption.

In the absence of peer pressure, we have a symmetric equilibrium with  $x_i^* = 79.61$  for all  $i \in N$ . The socially optimal consumption is  $x_i^0 = 31.19$  for all  $i \in N$ . So the electricity consumption is more than two and a half times the socially optimal level. The ratio of the marginal externality on any agent to the curvature of the raw utility of any other agent at the equilibrium consumption  $x^*$  is  $1.33 > c$  and at socially optimum consumption  $x^0$  is  $0.27 < c$ . Therefore the cost of exerting the socially optimal level of pressure on the peers is much higher than the resulting reduction in externality experienced by any agent. The symmetric equilibrium in the model with peer-pressure is better than the model without peer-pressure. The peer-pressure on any agent in equilibrium is 0.13, significantly lower than the optimal peer-pressure 0.83. Therefore, the consumption in equilibrium is 67.46 which is more than twice the socially optimal consumption. Figure 2 illustrates the situation, highlighting that even if peer-pressure is possible, it is too costly to apply sufficiently enough to yield socially optimal consumption.

The total reward budget required to reduce the consumption to the socially optimal level using the Pigouvian mechanism will be 3995.40. On the other hand, the total reward budget using the social mechanism will be 1095.26. Figure 3 shows that the required budget for any target consumption level is lower under the social mechanism than under the Pigouvian mechanism.

The social cost in equilibrium under the social mechanism will be 82.50. Moreover, the combined total social cost and the redistribution cost under the social mechanism is 192.03. This is half of the total redistribution cost of 399.54 under the Pigouvian mechanism.



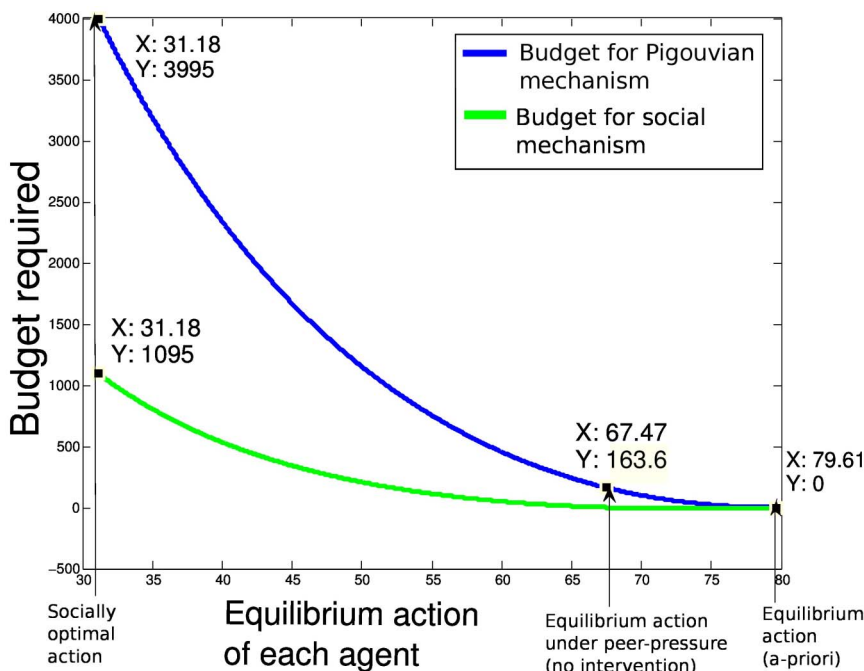
**Figure 2 | (navy) Raw utility of consumption, maximized at the equilibrium consumption of  $X = 79.62$ .** The peer pressure will lower the equilibrium consumption only slightly to  $X = 67.46$  due to high marginal cost of pressure; (green) Externality experienced due to other agents' consumption; (red) Total utility curve is the difference between the navy and green curves, and is maximized at the (much lower) socially optimal consumption level of  $X = 31.19$ .

Figure 4 shows that the redistribution and social loss for any target consumption level is lower under the social mechanism than under the Pigouvian mechanism.

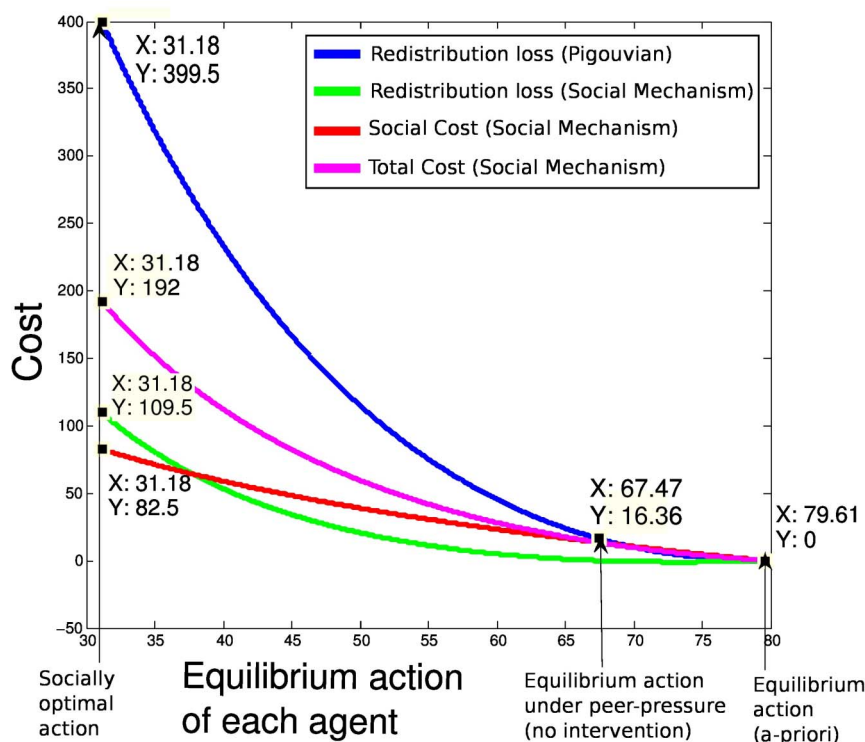
**Illustrative experiment: promoting physical activity.** We now discuss an experiment designed to verify the testable predictions of our results. This experiment was conducted by one of the authors (Pentland) and our colleagues, as part of the “Friends and Family”

study at MIT<sup>20</sup>. In particular, the social mechanism was tested against the Pigouvian mechanism. The details of the experimental set-up are provided in Ref. 20. Here, we analyze the results further, and interpret the findings within the theoretical framework of the present paper, providing empirical validation for it.

Our model and results are consistent with data observed in the experiment and the three testable predictions are verified. In the MIT “Friends and Family” study, the aim was to encourage people to



**Figure 3 | (navy) Budget required for rewards under Pigouvian mechanism for a target consumption level; (green) Budget required for rewards under Social mechanism for a desired consumption level.** It increases much slower as the target consumption level is reduced.



**Figure 4** | (navy) Redistribution loss under Pigouvian mechanism for a target consumption level; (green) Redistribution loss under Social mechanism for a desired consumption level. It increases much slower as the target consumption level is reduced; (red) Social loss under Social mechanism for a desired consumption level; (magenta) Total (redistribution + social) loss under Social mechanism for a desired consumption level.

increase their daily physical activity, as measured by the accelerometer sensors in their mobile phones. The relationships between individuals were captured through surveys and physical proximity measured by bluetooth data from the phones. Each individual was assigned two peers. Some individuals had close friends as their peers and some had strangers as peers. Three treatments were experimented with. In the self-monitoring (control) condition, individuals were rewarded for their own increased physical activity. This condition simulates the Pigouvian mechanism. Two treatment conditions involved assigning two “buddies” to each participant. In the “peer-view” treatment condition, the participant was shown the buddies’ activity levels, but was still rewarded for his/her own activity level. This condition simulates the game with peer-pressure under the Pigouvian mechanism. In the “peer-reward” condition, it was the buddies who received a reward proportional to the participant’s activity. This condition simulates the social mechanism. The peer-view condition mediates normative forces<sup>21</sup>, while the peer-reward condition was designed to promote peer pressure.

Before the mechanisms were deployed, the activity levels of the individuals were monitored for a period of 23 days. The average normalized activity levels (measured by accelerometers) of the three groups were as follows: (a) 1.162 for the control group, (b) 1.266 for the peer-see group, and (c) 1.216 for the peer-reward group. The marginal reward was computed such that the budget for the reward in all the three conditions was roughly the same. The marginal rewards were (a) \$ 83.3 (=1/.012) in the control condition, (b) \$ 39.5 (=1/.0253) in the peer-see condition (roughly half of the marginal reward in the control condition) and (c) \$ 12 (=1/(2 × 0.0416)), in the peer-reward condition (roughly one third the marginal reward in the peer-see condition) since the reward was given to two peers for each agent. After the mechanisms were deployed and removing the initial 19 days left for stabilizing the behaviors, the activity levels were monitored for the subsequent 20 days. The change in the average normalized activity levels (measured by accelerometers) of the three groups were as follows: (a) 0.037 (3.2%) for

the control group, (b) 0.070(5.5%) for the peer-see group, and (c) 0.126(10.4%) for the peer-reward group.

Results reveal that peer pressure is indeed present when the information about an individual’s actions is known to the peers. The “peer-see” condition showed a significant increase in activity level compared to the control (0.070 vs 0.037,  $p < 0.01$ ), even when the peers did not receive any reward, suggesting an effect of peer pressure. This verifies the first testable prediction.

The data reveals that at low marginal rewards, the peer-based mechanism yields higher returns. Marginal reward is the reward distributed per unit increase in physical activity. Despite lower marginal reward, the peer-reward condition yielded significantly larger increase in activity level than both the control and peer-see conditions (0.126 vs 0.07 and 0.037 respectively). This verifies the second testable prediction.

Furthermore, when both assigned peers of an individual were close friends, the peer-reward condition exhibited twice the improvement in the normalized activity levels as compared to the same condition with weaker social ties with the assigned peers. The change in the average normalized activity level for the peer-reward mechanism (a) in the triads consisting of close friends was 0.269 for an average reward of \$3.00 and (b) in the triads consisting of strangers was 0.137 for an average reward of \$2.95. This shows that when social capital is high, leading to lower cost of exerting pressure, the peer-reward mechanism has even better returns. This verifies the third testable prediction.

## Discussion

There is growing evidence that peer pressure can promote cooperative behavior. In evolutionary biology and behavioral economics, peer pressure, embodied in the form of costly punishment or reward, is recognized as a mechanism for promoting cooperation<sup>7,8</sup>. Various observations have also been anecdotally crediting peer pressure for improved cooperation in micro-loan payments<sup>11</sup> and natural





resource conservation<sup>14</sup>. Our main contribution is to place these results in a theoretical foundation in the theory of externalities. This enabled us to identify precisely how to design policies that explicitly employ peer pressure in regulating behavior to promote cooperation at a large-scale. Further, our analysis revealed precise conditions under which such policies outperform traditional centralized, Pigouvian policies.

A key advantage of our approach is that the precise means by which pressure is exerted is left to the individuals in the network, and thus works at the normative level<sup>31</sup>. To quote Nobel Prize winner Elinor Ostrom: “*International donors and nongovernmental organizations, as well as national governments and charities, have often acted, under the banner of environmental conservation, in a way that has unwittingly destroyed the very social capital – shared relationship, norms, knowledge and understanding – that has been used by resource users to sustain the productivity of natural capital over the ages*”<sup>22</sup>. Building on Ostrom’s sentiment, by adjusting the incentives at the local level, our peer-based incentive mechanism leverages existing social capital and local regulatory mechanisms, rather than bypassing (or even suppressing) them.

We note that our results must be taken with some caveats. Firstly, negative peer pressure may result in retaliatory action, especially when it is perceived as illegitimate<sup>29</sup>, a phenomenon that has been observed across many cultures<sup>30</sup>. Thus, peer enforcement can turn prohibitively expensive if peer pressure ends in a chain reaction of retaliation. Given this challenge, encouraging positive interpersonal pressure can be more effective, as it is less likely to result in retaliation<sup>7</sup>.

A second caveat is that the success of peer pressure relies on effective monitoring of peer action. In our model, the cost of monitoring is incorporated implicitly into the cost of pressure. However, comparing the monitoring costs of social and Pigouvian approaches is a very relevant question. One can imagine situations in which peer monitoring is very efficient compared to centralized solutions (e.g. neighbor monitoring of compliance with frontyard water usage restrictions in some countries). Yet one can also imagine domains in which peer monitoring may be too invasive or may require costly information acquisition.

In a recent review of data-centric uses of social networks to accelerate behavior change and improve collective performance, Valente identified four major classes of *network intervention*<sup>23</sup>: (i) targeting influential individuals to use them as conduits for desirable behavior; (ii) segmentation and targeting of groups with particular network characteristics; (iii) altering (or designing) the network structure; and (iv) induction of particular peer-to-peer interaction. The theoretical underpinnings of the first three intervention classes have received much attention: influential individual identification<sup>24–26</sup>, characteristic segmentation of groups<sup>27</sup>, and network structure design<sup>28</sup>. This paper fills an important gap by putting the induction of peer-to-peer influence on solid theoretical foundation.

## Methods

As mentioned above, our analysis is based on a two-stage game model of peer pressure and externality-causing action. In the first stage, the agents in the network choose the amount of peer pressure to exert on each peer. In the second stage, the agent choose their externality-causing action (e.g. energy consumption). Then, we show that under natural assumptions on the utility function (strictly concave and has a finite maximum) and the externality function (strictly convex and increasing), there exists at least one sub-game perfect equilibrium of the game. We characterize these sub-game perfect equilibria, using the Karush-Kuhn-Tucker (KKT) conditions, and show that the actions taken by any agent in the second stage of the game (i.e. their consumption action) is identical in all such equilibria. This implies that the pressure felt by any agent in the first stage of the game is identical in all equilibria.

The above analysis suggests that the only difference between the different equilibria is who puts pressure on whom, always yielding the same total pressure on any agent. Using the KKT analysis, under assumptions 1 and 2, we show that when agents have the ability to exert peer pressure, then the total consumption in equilibrium will be strictly lower.

Next, we show that under assumption 3, using the Pigouvian mechanism to internalize externalities, the equilibrium is identified by a boundary condition, in

which the cost of exerting peer pressure is too high compared to the marginal effect of peer pressure. Thus, there is no peer pressure at all under the Pigouvian mechanism.

We then introduce a simple linear social reward function. We show that the marginal social reward depends upon both the agent causing the externality, and the peers who can exert peer pressure on him. The reward function essentially compensates the peers for the cost of exerting pressure, while discounting for the reduction in externality that they benefit from.

Next, we find that the budget required in the social mechanism is lower than that of the Pigouvian mechanism. This is due to the amplification of the effect of rewards. This relies on the marginal effect of peer pressure being higher than the marginal effect of direct rewards to agents.

For our final result, we incorporate an additional quantity into the model, namely the *redistribution factor*, which captures the overhead involved in the implementation of any mechanism as a proportion of the total reward expended by the mechanism. This allows us to compare the Pigouvian mechanism, in which a larger reward is redistributed with a given redistribution loss, with the social mechanism, in which the redistribution overhead is lower, but it entails *social cost* to the agents exerting the peer pressure. For sparse networks, we show that the redistribution loss and social cost combined are still lower than the Pigouvian mechanism.

Details of our methods, along with proofs of all theoretical results, can be found in the Supporting Information appendix.

- Dietz, T., Ostrom, E. & Stern, P. The struggle to govern the commons. *Science* **302** (5652), 1907–1912 (2003).
- Hardin, G. Tragedy of the commons. *Science* **162** (3859), 1243–1248 (1968).
- Baumol, W. J. On taxation and the control of externalities. *American Economic Review* **62** (3) (1972).
- Coase, R. The problem of social cost. *Journal of Law and Economics* **3**, 1–44 (1960).
- Slemrod, J. Optimal taxation and optimal tax systems. *The Journal of Economic Perspectives* **4** (1), 157–178 (1990).
- Nowak, M. Five rules for the evolution of cooperation. *Science* **314** (5805), 1560–1563 (2006).
- Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M. A. Positive interactions promote public cooperation. *Science* **325**, 1272–1275 (2009).
- Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
- Calvo-Armengol, A. & Jackson, M. Peer pressure. *Journal of the European Economic Association* **8** (1), 62–89 (2010).
- Kandel, E. & Lazear, E. Peer pressure and partnerships. *Journal of Political Economy* **100** (4), 801–817 (1992).
- Karlan, D. Social connections and group banking. *The Economic Journal* **117** (517), F52–F84 (2007).
- Breza, E. Peer effects and loan repayment: Evidence from the krishna default crisis. *Job Market Paper MIT* (2012).
- Van Vugt, M. Community identification moderating the impact of financial incentives in a natural social dilemma: Water conservation. *Personality and Social Psychology Bulletin* **27** (11), 1440–1449 (2001).
- Van Vugt, M. Averting the tragedy of the commons using social psychological science to protect the environment. *Current Directions in Psychological Science* **18** (3), 169–173 (2009).
- Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466** (7308), 861–863 (2010).
- Mas-Colell, A., Whinston, M. D. & Green, J. R. *Microeconomic Theory*. Oxford University Press, (1995).
- Osborne, M. J. & Rubinstein, A. *A course in game theory*. MIT Press, (1994).
- Sobel, J. Can we trust social capital? *Journal of economic literature* **40** (1), 139–154 (2002).
- Woolcock, M. & Narayan, D. Social capital: Implications for development theory, research, and policy. *The world bank research observer* **15** (2), 225–249 (2000).
- Aharony, N., Pan, W., Ip, C., Khayal, I. & Pentland, A. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing* **7**, 643–659 (2011).
- Schultz, P., Nolan, J., Cialdini, R., Goldstein, N. & Griskevicius, V. The constructive, destructive, and reconstructive power of social norms. *Psychological Science* **18** (5), 429–434 (2007).
- Ostrom, V. & Ostrom, E. (November 7, 2003). Rethinking Institutional Analysis: Interview at the Mercatus Center, George Mason University..
- Valente, T. W. Network interventions. *Science* **337** (6090), 49–53 (2012).
- Kempe, D., Kleinberg, J. & Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146. ACM, (2003).
- Liu, Y., Slotine, J. & Barabási, A. Controllability of complex networks. *Nature* **473** (7346), 167–173 (2011).
- Aral, S. & Walker, D. Identifying influential and susceptible members of social networks. *Science* **337** (6092), 337–341 (2012).
- Newman, M. & Girvan, M. Finding and evaluating community structure in networks. *Physical review E* **69** (2), 026113 (2004).
- Johnson, D., Lenstra, J. & Kan, A. The complexity of the network design problem. *Networks* **8** (4), 279–285 (2006).





29. Nikiporakis, N. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* **92** (1), 91–112 (2008).
30. Herrmann, B., Thöni, C. & Gächter, S. Antisocial punishment across societies. *Science* **319** (5868), 1362–1367 (2008).
31. Kandori, M. Social norms and community enforcement. *The Review of Economic Studies*. **59** (1), 63–80 (1992).

## Acknowledgments

This work was partially funded by the Martin Family Fellowship for Sustainability. This research was partially sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053, by AFOSR under Award Number FA9550-10-1-0122, and by Masdar Institute of Science and Technology. Views and conclusions in this document are those of the authors and should not be interpreted as representing the policies, either expressed or implied, of the sponsors.

## Author contributions

A.M., I.R. and A.P. performed analyses, discussed the results, and contributed to the text of the manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

**How to cite this article:** Mani, A., Rahwan, I. & Pentland, A. Inducing Peer Pressure to Promote Cooperation. *Sci. Rep.* **3**, 1735; DOI:10.1038/srep01735 (2013).