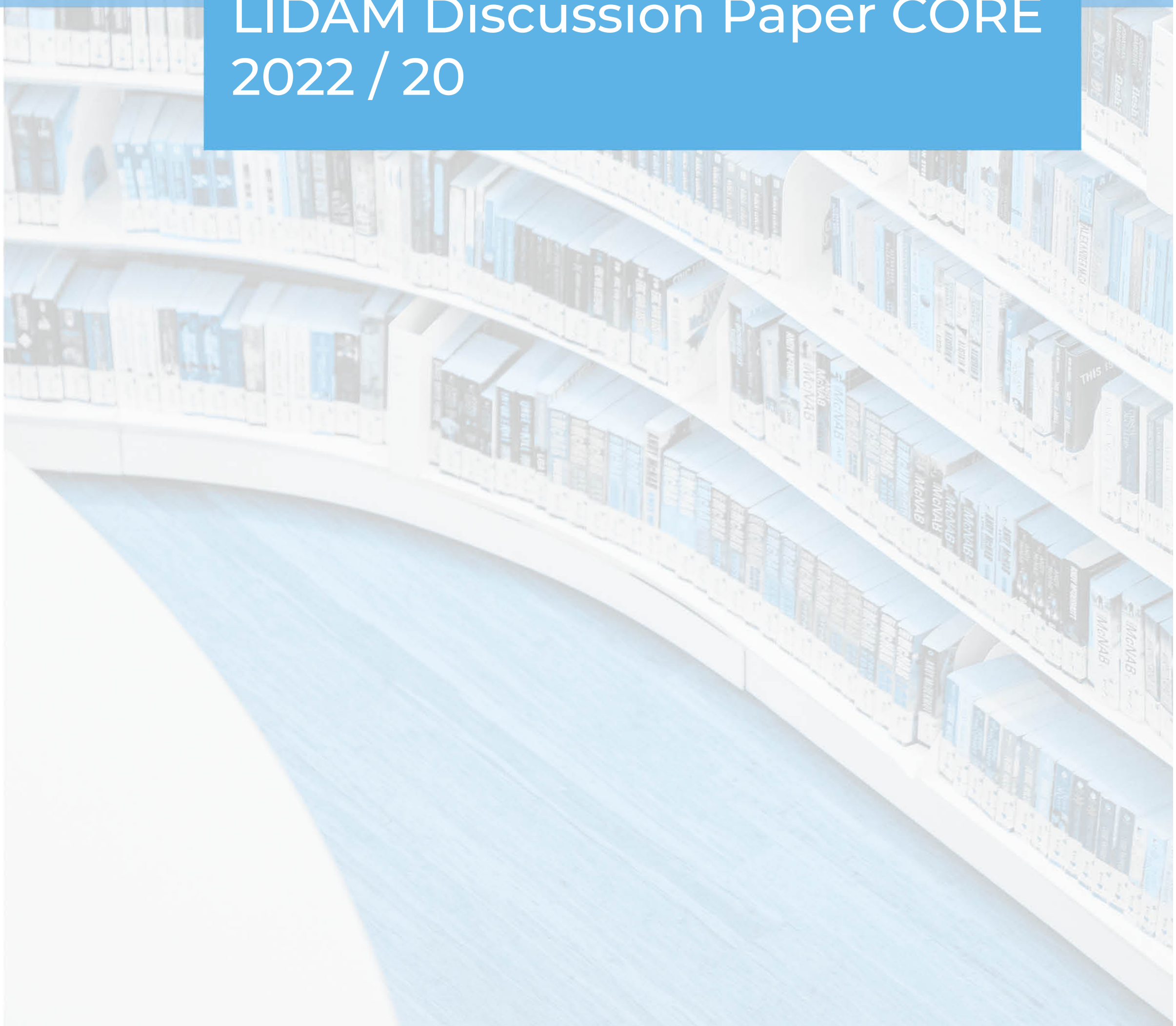
KEY PLAYERS IN BULLYING NETWORKS

Ata Atay, Ana Mauleon, Simon Schopohl, Vincent Vannetelbosch



LIDAM Discussion Paper CORE
2022 / 20

UCLouvain

LIDAM
CORE | Louvain Institute of Data Analysis and Modeling in economics and statistics

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers.html

Key players in bullying networks

Ata Atay* Ana Mauleon† Simon Schopohl‡
Vincent Vannetelbosch§

May 23, 2022

Abstract

Individuals are embedded in a network of relationships and they can be victims, bystanders, or perpetrators of bullying and harassment. Each individual decides non-cooperatively how much effort to exert in preventing misbehavior. Each individual's optimal effort depends on the contextual effect, the social multiplier effect and the social conformity effect. We characterize the Nash equilibrium and we derive an inter-centrality measure for finding the key player who once isolated increases the most the aggregate effort. An individual is more likely to be the key player if she is influencing many other individuals, she is exerting a low effort because of her characteristics, and her neighbors are strongly influenced by her. The key player policy increases substantially the aggregate effort and the targeted player should never be selected randomly. The key player is likely to remain the key player in presence of social workers except if she is becoming much less influential due to her closeness to social workers. Finally, we consider alternative policies (e.g. training bystanders for helping victims) and compare them to the policy of isolating the key player.

Keywords: Social networks · bullying · harassment · peer effects · key player · conformity · #MeToo
JEL Classifications: A14 · C72 · D85 · Z13

*Department of Mathematical Economics, Finance and Actuarial Sciences, and Barcelona Economic Analysis Team (BEAT), University of Barcelona, Spain (aatay@ub.edu)

†CEREC, UCLouvain Saint-Louis Brussels and CORE/LIDAM, UCLouvain, Belgium (ana.mauleon@usaintlouis.be)

‡CEREC, UCLouvain Saint-Louis Brussels, Belgium

§CORE/LIDAM, UCLouvain, Belgium (vincent.vannetelbosch@uclouvain.be)

1 Introduction

More than one out of every five students report being bullied in the United States (National Center for Educational Statistics, 2016). Rates of bullying for 12-18 year old students are around 35% for traditional bullying involvement and 15% for cyberbullying involvement.[1] The reasons for being bullied reported most often by students are related to students' characteristics like race/ethnicity, gender, disability, religion, sexual orientation. Bullying has a negative effect on the physical and mental health of bullied students, on their school work, on their relationships with friends and family, and on how they feel about themselves. Hence, it is important to implement successful strategies to prevent bullying.

Hawkins, Pepler and Craig (2001) find that more than half of bullying situations stop when peers or friends intervene on behalf of the student being bullied. In addition, the decision of one bystander to exert more effort in reporting and helping bullied students positively influence the behavior of her peers. This is the so-called social multiplier effect.[2] One type of interventions implemented to reduce the negative effects of bullying are peer norms interventions that make clear to young people that most of their peers oppose mistreatment.[3] Once bystanders are influenced by peer norms, it becomes even more costly for them to exert less effort than their peers in reporting and helping bullied students. This is the so-called social conformity effect.

In this paper, we propose the key player strategy to reduce bullying and harassment. The key player strategy aims at finding and isolating optimally a negatively influencing individual in order to increase the reporting of bullying and harassment as well as effective prevention and intervention efforts by peers within the social network.[4]

We adopt an approach similar to Ballester, Calvo-Armengol and Zenou (2006) and Ballester and Zenou (2014) for identifying key players. We develop a network game[5] where individuals are connected through a network and they can be victims, bystanders, perpetrators or social workers. Individuals decide non-cooperatively how much effort to exert in reporting bullying or harassment. The individual effort to report misbehavior may

[1]See Modecki, Minchin, Harbaugh, Guerra and Runions (2014). Of these students who report being bullied at school, 33% indicate that they are bullied at least once or twice a month during the school year. A slightly higher portion of female than of male students report being bullied at school. But, a higher percentage of male than of female students report being physically bullied (National Center for Educational Statistics, 2016).

[2]Thornberg, Tenenbaum, Varjas, Meyers, Jungert and Vanegas (2012) find that beliefs of bystanders in their social self-efficacy are positively associated with defending and negatively associated with passive behavior from other bystanders.

[3]See Davis and Nixon (2013). Other interventions often implemented are building staff-students connections, disciplinary responses to negative peer behaviors, encouraging bystanders to confront and discourage the unkind behavior, and social skills training.

[4]Mouttapa, Valente, Gallaher, Rohrback, and Unger (2004) found the friendship network and the pattern of friendships among individuals within a group are important aspects of adolescent school bullying.

[5]Jackson and Zenou (2015) provide a comprehensive introduction to network games.

be affected by (i) her individual characteristics and the characteristics of her neighbors (i.e. the contextual effect), (ii) the effort levels of her neighbors (i.e. the social multiplier or network spillovers effect), (iii) the norms of conduct set by neighbors[6] (i.e. the social conformity effect), and (iv) unobservable correlated effects. Social workers are assumed to exert more effort than what they would optimally do if they were standard individuals.

We derive the unique Nash equilibrium of the network game. The equilibrium effort of each individual is proportional to her Katz-Bonacich weighted centrality. We look for the key player. The key player is defined as the individual who once isolated generates the greatest augmentation in the total effort for reporting bullying or harassment. To do so we propose a new measure of contextual inter-centrality that determines the key player to be isolated. This measure captures three effects: (i) the change in effort due to the change in the context when some individual is isolated, (ii) the change in effort due to the network structure change after the isolation, (iii) the effort exerted by the individual who is isolated.[7]

We illustrate the policy of finding and isolating the key player by means of a specific network structure that is rich enough to disentangle the effects of network spillovers, social conformity and players' characteristics on the resulting equilibrium outcomes and the key player strategy. We find that the key player is not necessarily the individual who is the most central within the network nor the individual who is doing less effort than all other individuals. The most central individual with a negative attribute (i.e. someone who could be a perpetrator) is not necessarily the key player to be isolated. In fact, an individual is more likely to become the key player if (i) she is influencing many other individuals (i.e. she has many neighbors), (ii) she is exerting a low effort because of her characteristics, and (iii) her neighbors are strongly influenced by her (i.e. her neighbors have few links). Implementing the key player strategy always increases the total effort exerted by all individuals except when individuals are homogeneous or their characteristics are correlated with their centrality. Comparing the total effort obtained by isolating the key player to the total effort that would be exerted if the target was selected randomly, we observe that the key player policy increases substantially the total effort. Hence, the planner should never target randomly the individual to be isolated. In the presence of social workers, an individual is more likely to become the key player if she is influencing negatively many other players and she is not too influenced by social workers. Hence,

[6]That is, individuals are penalized if they deviate from the effort level of their neighbors in the social network. See e.g. Patacchini and Zenou (2012), Boucher (2016), Landini, Montinari, Pin and Piovesan (2016), Boguslaw (2017), Lee, Liu, Patacchini and Zenou (2021) about peer effects and conformism in social networks.

[7]Only the first two effects are present in the contextual inter-centrality measure of Ballester and Zenou (2014) where the key player is removed rather than being isolated and the planner's objective is to reduce the total effort level. In our context, the objective of the planner is to increase the total effort in reporting misbehavior and an isolated individual may still exert some positive effort depending on her own characteristics. Hence, the effort of the isolated individual matters for determining who is the key player.

an individual who was the key player without social workers is likely to remain the key player except if she is becoming less central and less influential due to her closeness to social workers.

We also consider alternative policies and we compare them to the policy of isolating the key player. A first alternative policy consists of finding the key player who once turned into a social worker generates the highest possible increase in aggregate effort level. To do so, we obtain the benevolent change inter-centrality measure. A second alternative policy consists of finding the key player who once trained for helping victims and reporting misbehavior generates the highest possible increase in aggregate effort level. Training some targeted individual modifies her characteristics so that she is now eager to exert more effort for reporting bullying and harassment. The planner has always incentives to implement both policies instead of doing nothing. However, both policies perform only slightly better than selecting randomly some individual. Thus, if the data collection about the relationships and the characteristics of the individuals is too costly, the planner might prefer to target randomly some individual who could be either turned into a social worker or trained for helping victims instead of implementing the policy of isolating the key player.

One of the first application of the key player strategy was developed for delinquent networks. Ballester, Calvo-Armengol and Zenou (2010) propose a delinquent network game where players decide about how much effort to exert in criminal activities. They derive both the key player (i.e. optimal single player removal for reducing criminal activities) and the key group (i.e. optimal group removal). Zenou (2016) gives an overview of the recent literature on key players in social and economic networks. There is an empirical literature that support key player policies. Using data from adolescents in the United States, Lee, Liu, Patacchini and Zenou (2021) show that contextual effects matter since the key player in crime may be different when one uses either Ballester, Calvo-Armengol and Zenou (2006) inter-centrality measure or Ballester and Zenou (2014) contextual inter-centrality measure. Moreover, compared to a policy that removes the most active delinquent from the network, they show that the key player strategy leads to a much higher delinquency reduction. Similarly, using a data set of co-offenders in Sweden, Lindquist and Zenou (2019) find that the key player strategy outperforms alternative policies like targeting the most active delinquent or targeting the most central delinquents in the criminal network.

Compared to other applications of the key player strategy, the present paper has the following innovations: (i) the key player is the individual who, once isolated from the rest of individuals, increases the most the aggregate effort exerted in the whole population, (ii) the key player is isolated rather than being removed from the network, (iii) individuals can be victims, bystanders, perpetrators or social workers, (iv) alternative key player strategies such as turning individuals into social workers or training individuals for helping victims are considered.

Our model can also be used to develop network-orientated strategies for increasing the reporting of sexual and gender-based violence (SGBV).[8] SGBV in close relationships is a widespread phenomenon found in societies all over the world. Almost one out of three women who have been in a relationship is estimated to have been abused by a partner during her lifetime (WHO, 2021).[9] Victims often do not report the violence or harassment they suffer. Shame, a desire to protect the perpetrator, stigma, guilt and fear are the main reasons that women subjected to SGBV give when explaining why they barely report SGBV. The abuse often concerns various aspects of women's everyday lives, affecting their social and economic situations, but also their physical and mental health (Boethius and Åkerström, 2020). Social network ties may try to stop the abuse, may help the victim by offering means of escape, or they may help report the violence. In October 2017 the actress Alyssa Milano used the hashtag #MeToo in a Twitter post: "if all the women who have been sexually harassed or assaulted wrote 'Me too' as a status, we might give people a sense of the magnitude of the problem." A viral campaign started and many men and women used the hashtag #MeToo to report harassment (social-multiplier effect). The perpetrators had to face serious consequences, while the victims were encouraged by others to use the hashtag #MeToo and report (social-conformity effect).

The paper is organized as follows. In Section 2 we present the bullying network game and we determine the Nash equilibrium effort levels of this game. In Section 3 we derive the contextual inter-centrality measure for finding the key player to be isolated and we study the relative performance of this key player policy. In Section 4 we consider alternative key player policies where the key player is either turned into a social worker or trained for helping victims. Finally, we draw conclusions.

2 Bullying network outcomes

2.1 The bullying network game

Let $N = \{1, \ldots, n\}$ be the finite set of rational players, S be the finite set of benevolent players or social workers, and $N^+ = N \cup S$ denote the set of all players with $\#N = n$ and $\#S = s$. Players are arranged in a network g where a link between player i and player j is denoted by $g_{ij} = 1$. If i and j are not linked, then $g_{ij} = 0$. By convention, $g_{ii} = 0$. Let $N_i = \{j \in N^+ \mid g_{ij} = 1\}$ be the set of neighbors of player i in g. We keep track of social connections in network through the row-normalized adjacency matrix $G^* = \left(g^*_{ij}\right)$. It is a directed and weighted network with $g^*_{ij} = g_{ij}/\sum_{j=1}^{n+s} g_{ij}$, and so for each player $i \in N^+$, $\sum_{j\in N^+} g^*_{ij} = 1$. If there is a link between i and j, then $g^*_{ij} > 0$. Otherwise, $g^*_{ij} = 0$.

[8] Based on our model, Ogbe, Jbour, Rahbari, Unnithan and Degomme (2021) analyze the potential impact of alternative network-oriented interventions for survivors of SGBV among asylum seekers in Belgium.

[9] WHO report 'Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence.' Department of Reproductive Health and Research (https://www.who.int/publications/i/item/9789241564625)

Consider some social network g where each player is potentially victim of bullying or harassment. Players in the network decide how much effort to exert in reporting bullies and/or in helping victims of bullying. We denote by x_i the effort level of player i, with $0 \leq x_i \leq \bar{x}$, and by $\mathbf{x} = (x_1, \dots, x_{n+s})$ the population profile. Each rational player $i \in N$ exerts an effort level that maximizes her utility:

$$u_i(x, G^*) = \phi_i \cdot x_i + \lambda_1 \cdot \left(\sum_{j \in N^+} g^*_{ij} \cdot x_j \right) \cdot x_i - \frac{\lambda_2}{2} \cdot \sum_{j \in N^+} g^*_{ij} \cdot (x_i - x_j)^2 - \frac{\lambda_3}{2} \cdot x_i^2$$

where

$$\phi_i = \mathbf{y}'_i \cdot \beta_1 + \left(\sum_{j \in N^+} g^*_{ij} \cdot \mathbf{y}'_j \right) \cdot \beta_2 + \xi + \epsilon_i$$

is the contextual effect of player i. The first term of the utility function, $\phi_i \cdot x_i$, describes the direct benefit from reporting or helping. Player i's contextual effect, ϕ_i, depends not only on her own attribute (e.g., age, gender, education, ...) but also on the weighted average attribute of her neighbors. The vector $\mathbf{y}_i$ contains all observable characteristics of player i and is transformed into a positive real number by means of β_1, while $\sum_{j \in N^+} g^*_{ij} \cdot \mathbf{y}'_j$ captures the weighted average characteristics of her neighbors $j \in N_i$ and is transformed into a positive real number by means of β_2. Finally, ξ and ϵ_i are error terms.

The second term of the utility function, $\lambda_1 \cdot \left(\sum_{j \in N^+} g^*_{ij} \cdot x_j \right) \cdot x_i$, captures the positive spillover effect from the effort exerted by the neighbors, weighted by λ_1. The parameter $\lambda_1 \geq 0$ is the social multiplier coefficient and it captures the strength of social multiplier effect. The decision of one player to exert more effort in reporting can directly influence the behavior of her neighbors or peers. That is, from each neighbor j that exerts a positive effort, player i obtains a spillover according to the weight of the link between i and j that induces her to exert a higher effort.

The third term of the utility function, $(\lambda_2/2) \cdot \sum_{j \in N^+} g^*_{ij} \cdot (x_i - x_j)^2$, captures the social conformity effect. The parameter $\lambda_2 \geq 0$ is the social conformity coefficient and it captures the strength of social conformity. Players are influenced by the social norm, and so there is a cost for deviating from the social norm which is increasing with the distance from the effort levels done by peers. That is, each player would like that her effort matches with the effort of her peers. Hence, the further the effort of player i and her neighbor j are away from each other, the larger is the conflict. Again, the weight of the link g_{ij} is used as an indicator of how much player i cares about having a conflict with player j. Notice that the efforts of network neighbors are strategic complements:

$$\frac{\partial^2 u_i(x, G^*)}{\partial x_i \partial x_j} = (\lambda_1 + \lambda_2) \cdot g^*_{ij} \geq 0,$$

with $\lambda_1, \lambda_2 \geq 0$.

The final term of the utility function, $(\lambda_3/2) \cdot x_i^2$, is the direct cost to exert the effort x_i for reporting bullies and/or helping victims of bullying. The higher the effort x_i the higher the direct cost. From now on, we assume that $\overline{x} \geq \phi_i/(\lambda_3 - \lambda_1)$ holds for each player $i \in N$. This condition ensures that the equilibrium effort level of player i is within the interval $[0, \overline{x}]$.

If a rational player $i \in N$ has no neighbor at all, she chooses an effort x_i, with $0 \leq x_i \leq \bar{x}$ to simply maximize

$$u_i(x_i) = (\mathbf{y}_i' \cdot \beta_1 + \xi + \epsilon_i)\, x_i - \frac{\lambda_3}{2} \cdot x_i^2.$$

Benevolent players or social workers are the ones who aim that bullies are reported and victims are helped. Thus, each benevolent player or social worker $i \in S$ always exerts an exogenous large effort x_S with $0 < x_S \leq \bar{x}$.

2.2 Nash Equilibrium

We first derive the optimal effort level of each rational player $i \in N$.

Lemma 1. *The best response function of a rational player $i \in N$ is given by*

$$x_i = \begin{cases} \max\{0, (\mathbf{y}_i' \cdot \beta_1 + \xi + \epsilon_i)\,/\lambda_3\} & \text{if} \quad N_i = \emptyset \\ \alpha_i + \gamma \cdot \sum_{j \in N} g_{ij}^* x_j & \text{if} \quad N_i \cap N \neq \emptyset \\ \alpha_i & \text{if} \quad \emptyset \neq N_i \subseteq S \end{cases} \tag{1}$$

with

$$\gamma = \frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3} \quad \text{and}$$

$$\alpha_i = \frac{\phi_i}{\lambda_2 + \lambda_3} + \gamma \cdot \sum_{j \in S} g_{ij}^* x_S.$$

Proof. We take $\partial u_i(x_i)/\partial x_i = 0$ and solve it for x_i. This is straightforward for i such that $N_i = \emptyset$. For i such that $N_i \cap N \neq \emptyset$ we get $\partial u_i(x_i)/\partial x_i = 0$

$$\begin{aligned}
\Leftrightarrow \quad & \phi_i + \lambda_1 \cdot \sum_{j \in N^+} g_{ij}^* x_j - \frac{1}{2} \cdot \lambda_2 \cdot \sum_{j \in N^+} g_{ij}^* \cdot (2x_i - 2x_j) - \frac{1}{2} \cdot \lambda_3 \cdot 2x_i &= 0 \\
\Leftrightarrow \quad & \phi_i + (\lambda_1 + \lambda_2) \cdot \sum_{j \in N^+} g_{ij}^* x_j - (\lambda_2 + \lambda_3) \cdot x_i &= 0 \\
\Leftrightarrow \quad & \frac{\phi_i}{\lambda_2 + \lambda_3} + \frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3} \cdot \sum_{j \in N^+} g_{ij}^* x_j &= x_i \\
\Leftrightarrow \quad & \frac{\phi_i}{\lambda_2 + \lambda_3} + \frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3} \cdot \sum_{j \in S} g_{ij}^* \underbrace{x_j}_{= x_S} + \frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3} \cdot \sum_{j \in N} g_{ij}^* x_j &= x_i \\
\Leftrightarrow \quad & \underbrace{\frac{\phi_i}{\lambda_2 + \lambda_3} + \frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3} \cdot \sum_{j \in S} g_{ij}^* x_S}_{\equiv \alpha_i} + \underbrace{\frac{\lambda_1 + \lambda_2}{\lambda_2 + \lambda_3}}_{\equiv \gamma} \cdot \sum_{j \in N} g_{ij}^* x_j &= x_i.
\end{aligned}$$

For i such that $\emptyset \neq N_i \subseteq S$, the second term of the last expression vanishes. □

Notice that the best response function of all connected rational players is independent of the effort of disconnected players. To find the Nash equilibria, we assume without loss of generality that there are no disconnected rational players.[10] Then, we get the $n \times 1$ vector $x = \alpha + \gamma G_n^* \cdot x$, where $\alpha = (\alpha_1, \dots \alpha_n)$ is a $n \times 1$ vector and G_n^* is the matrix G^* limited to the first n lines and n columns. Let $\mathbb{1}_n$ be the n-dimensional vector of ones.

Proposition 1. *Assume that the spectral radius ρ of G_n^* satisfies $\rho(G_n^*) \cdot \gamma < 1$. The unique Nash equilibrium in pure strategies is given by*

$$x^* = (\mathbb{1}_n - \gamma G_n^*)^{-1} \cdot \alpha.$$

Proof. The unique Nash equilibrium is obtained by solving the best response for x. From Lemma 1, we have

$$\begin{aligned} x^* &= \alpha + \gamma G_n^* x^* \\ x^* - \gamma G_n^* x^* &= \alpha \\ (\mathbb{1}_n - \gamma G_n^*) \cdot x^* &= \alpha \\ x^* &= (\mathbb{1}_n - \gamma G_n^*)^{-1} \cdot \alpha \end{aligned}$$

We require that $\rho(G_n^*) \cdot \gamma < 1$ to ensure that $(\mathbb{1}_n - \gamma G_n^*)^{-1}$ is well-defined and non-negative (Debreu and Herstein, 1953). □

Let $(G_n^*)^k$ be the k-th power of G_n^*, with coefficients $g_{ij}^{[k]}$, where k is some integer. The matrix $(G_n^*)^k$ keeps track of the weighted indirect connections in the network: $g_{ij}^{[k]} \geq 0$ measures the weight of walks of length $k \geq 1$ from i to j that go through only rational players. Given a scalar $\gamma \geq 0$ and G_n^*, we define the following matrix:

$$M = (\mathbb{1}_n - \gamma G_n^*)^{-1} = \sum_{k=0}^{\infty} \gamma^k (G_n^*)^k.$$

Given a $n \times 1$ vector α, we define the Katz-Bonacich α-weighted centrality (due to Bonacich, 1987) of parameter γ as

$$\mathbf{b}(g^*, \alpha, \gamma) = \sum_{k=0}^{\infty} \gamma^k (G_n^*)^k \alpha = (\mathbb{1}_n - \gamma G_n^*)^{-1} \cdot \alpha.$$

Corollary 1. *Assume that the spectral radius ρ of G_n^* satisfies $\rho(G_n^*) \cdot \gamma < 1$. Then, the unique Nash equilibrium in pure strategies is given by*

$$x^* = \mathbf{b}(g^*, \alpha, \gamma).$$

[10] If there are k disconnected players, we simply remove the entries corresponding to the disconnected players from the vector α and the matrix G_n^* and we obtain α as a $n-k \times 1$ vector and G_n^* as a $n-k \times n-k$ matrix.

Thus, at the Nash equilibrium, each player $i \in N$ exerts an effort equal to her weighted Katz-Bonacich centrality.

Notice that, in the case players are homogeneous (i.e. all rational players have the same characteristics and no benevolent player), then players will exert the same equilibrium effort given by

$$x_i^* = \frac{\phi_i}{\lambda_3 - \lambda_1} \text{ for all } i \in N,$$

with $\phi_j = \phi_k$ for all $j, k \in N$. As expected, the social conformity coefficient λ_2 does not affect the equilibrium outcomes when all neighbors are identical in their attributes or characteristics. In addition, regardless of the network structure and the number of players connected, the equilibrium efforts of all players are identical and depend only on λ_1, λ_3 and ϕ. Hence, aggregate utilities and total efforts are the same for different network architectures connecting all homogeneous players. This result is driven by the fact that we keep track of social connections through the row-normalized weighted adjacency matrix.

3 Bullying network policies

3.1 Finding and isolating the key player

We denote the entries of M by m_{ij} and the entries of the $n \times 1$ vector $\mathbf{b}$ by b_k. It holds that

$$b_k(g^*, \alpha) = \sum_{j \in N} m_{jk} \alpha_j.$$

The planner's objective is to find the key player, that is, the rational player who once isolated generates the highest possible increase in aggregate effort level. Player i is said to be isolated if $N_i = \emptyset$. To find the key player, we have to compare the original network with the network where a player is removed. Let $G^{*[-i]}$ denotes the adjacency matrix in which player i has been removed from the network. This adjacency matrix $G^{*[-i]}$ is obtained from G^* by removing the ith row and the ith column and by adjusting the weights such that the weights of all outgoing links sum up to 1: $g_{jk}^{*[-i]} = g_{jk} / \sum_{k \in N^+ \backslash \{i\}} g_{jk}$. When player i is removed from the network, the matrix M becomes

$$M^{[-i]} = \left(\mathbb{1}_{n-1} - \gamma G_n^{*[-i]}\right)^{-1}$$

with $m_{jk}^{[-i]}$ being the entries of matrix $M^{[-i]}$. Let $c_k(g^*, \alpha) = \sum_{j \in N} m_{jk} \alpha_k$, where the index of α in the summation is the only difference between b_k and c_k, and let

$$\alpha_j^{[-i]} = \frac{\phi_j^{[-i]}}{\lambda_2 + \lambda_3} + \gamma \cdot \left(\sum_{k \in S} g_{jk}^{*[-i]} x_b \right)$$

$$\phi_j^{[-i]} = \mathbf{y}_j' \cdot \beta_1 + \left(\sum_{k \in N^+, k \neq i} g_{jk}^{*[-i]} \cdot \mathbf{y}_k' \right) \cdot \beta_2 + \xi \epsilon_j.$$

Let $B(g^*,\alpha)$ denote the total effort of all rational players. It holds that

$$B(g^*,\alpha) = \sum_{j\in N} b_j(g^*,\alpha) = \sum_{j\in N} c_j(g^*,\alpha).$$

The total effect of isolating player i is given by

$$B(g^{*[-i]},\alpha^{[-i]}) - B(g^*,\alpha) + \max\left\{0, \frac{\mathbf{y}_i'\cdot\beta_1+\xi+\epsilon_i}{\lambda_3}\right\} = -\delta_i^1$$

To find who is the key player to be isolated, we simply need to solve $\min_{i\in N}\delta_i^1$.

Proposition 2. *Assume that each player has at least two links. Then, the contribution of isolated player i to the total effort in the game is given by the contextual inter-centrality:*

$$\delta_i^1(g^*,\alpha) = \underbrace{\sum_{\substack{k\in N\\ k\in N_i}} c_k(g^{*[-i]},\alpha) - c_k(g^{*[-i]},\alpha^{[-i]})}_{\textit{Contextual change}} + \underbrace{\frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j\in N} m_{ij}}_{\textit{Intercentrality of player } i}$$
$$-\underbrace{\max\left\{0, \frac{\mathbf{y}_i'\cdot\beta_1+\xi+\epsilon_i}{\lambda_3}\right\}}_{\textit{Effort of } i \textit{ when isolated}}.$$

Proof. Under the assumption that each player has at least two links, the contribution of player i to the total effort in the game is

$$B(g^*,\alpha) - B(g^{*[-i]},\alpha^{[-i]}) + \max\left\{0, \frac{\mathbf{y}_i'\cdot\beta_1+\xi+\epsilon_i}{\lambda_3}\right\}.$$

We have that

$$B(g^*,\alpha) - B(g^{*[-i]},\alpha^{[-i]}) = \sum_{j=1}^{n}\sum_{k=1}^{n} m_{jk}\alpha_k - \sum_{j=1}^{n}\sum_{k=1}^{n} m_{jk}^{[-i]}\alpha_k^{[-i]}$$
$$= \sum_{j=1}^{n}\Bigg(\sum_{\substack{k=1\\k\neq i}}^{n} m_{jk}\alpha_k - m_{jk}^{[-i]}\alpha_k^{[-i]}\Bigg) + \sum_{j=1}^{n} m_{ji}\alpha_i$$
$$= \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n} m_{jk}\alpha_k - m_{jk}^{[-i]}\alpha_k^{[-i]} + \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i\\k\notin N_i}}^{n} m_{jk}\alpha_k - m_{jk}^{[-i]}\alpha_k^{[-i]} + \sum_{j=1}^{n} m_{ji}\alpha_i$$
$$= \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n} m_{jk}\alpha_k - m_{jk}^{[-i]}\alpha_k + m_{jk}^{[-i]}\alpha_k - m_{jk}^{[-i]}\alpha_k^{[-i]} + \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i\\k\notin N_i}}^{n} (m_{jk} - m_{jk}^{[-i]})\alpha_k + \sum_{j=1}^{n} m_{ji}\alpha_i$$
$$= \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n} m_{jk}^{[-i]}(\alpha_k - \alpha_k^{[-i]}) + \sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n} (m_{jk} - m_{jk}^{[-i]})\alpha_k + \sum_{j=1}^{n} m_{ji}\alpha_i$$

$$= \sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n}\sum_{j=1}^{n} m_{jk}^{[-i]}(\alpha_k - \alpha_k^{[-i]}) + \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij}$$

$$= \sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n} c_k(g^{*[-i]},\alpha) - c_k(g^{*[-i]},\alpha^{[-i]}) + \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij},$$

where the fourth equality follows from the fact that $\alpha_k = \alpha_k^{[-i]}$ for all $k \notin N_i$. The sixth equality follows from

$$\sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}(m_{jk} - m_{jk}^{[-i]})\alpha_k + \sum_{j=1}^{n} m_{ji}\alpha_i = \sum_{j=1}^{n} b_j(g^*,\alpha) - \sum_{\substack{j=1\\j\neq i}}^{n} b_j^{[-i]}(g^{*[-i]},\alpha)$$

$$= b_i(g^*,\alpha) + \sum_{\substack{j=1\\j\neq i}}^{n} b_j(g^*,\alpha) - b_j^{[-i]}(g^{*[-i]},\alpha) = b_i(g^*,\alpha) + \sum_{\substack{j=1\\j\neq i}}^{n}\sum_{k=1}^{n} m_{jk}\alpha_k - m_{jk}^{[-i]}\alpha_k$$

$$= b_i(g^*,\alpha) + \sum_{\substack{j=1\\j\neq i}}^{n}\sum_{k=1}^{n}\frac{m_{ij}m_{ik}}{m_{ii}}\alpha_k = b_i(g^*,\alpha) + b_i(g^*,\alpha)\cdot\sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}}{m_{ii}}$$

$$= b_i(g^*,\alpha)\Big(1 + \sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}}{m_{ii}}\Big) = b_i(g^*,\alpha)\Big(\frac{m_{ii}}{m_{ii}} + \sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}}{m_{ii}}\Big) = \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij},$$

where the fourth equality follows from Lemma 1 in Ballester, Calvo-Armengol and Zenou (2006) and the fifth equality is obtained from

$$\sum_{\substack{j=1\\j\neq i}}^{n}\sum_{k=1}^{n}\frac{m_{ij}m_{ik}}{m_{ii}}\alpha_k = \sum_{\substack{j=1\\j\neq i}}^{n}\sum_{k=i}\frac{m_{ij}m_{ik}}{m_{ii}}\alpha_k + \sum_{\substack{j=1\\j\neq i}}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}\frac{m_{ij}m_{ik}}{m_{ii}}\alpha_k$$

$$= \sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}m_{ii}}{m_{ii}}\alpha_i + \sum_{\substack{j=1\\j\neq i}}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}\frac{m_{ij}m_{ik}}{m_{ii}}\alpha_k = \sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}}{m_{ii}}\Big(m_{ii}\alpha_i + \sum_{k=1,k\neq i}^{n} m_{ik}\alpha_k\Big)$$

$$= b_i(g^*,\alpha)\cdot\sum_{\substack{j=1\\j\neq i}}^{n}\frac{m_{ij}}{m_{ii}}.$$

□

The decision of reporting of each individual is affected not only by her own characteristics but also by the characteristics of her friends. The contextual inter-centrality of player i, $\delta_i^1(g^*,\alpha)$, emphasizes the three effects at work when player i is isolated from the rest of the players. The first effect is the contextual variable change effect, which is due to the change in the context when player i is isolated from the network while the network is kept unchanged. The second effect is the network structure change effect, which captures the change in effort due to the network structure change after the removal of player i. The third effect is simply the effort exerted by player i when isolated. Only the first two effects are present in the contextual inter-centrality measure of Ballester and Zenou (2014) where the key player is removed rather than being isolated. However, an isolated player may still exert some positive effort in reporting and her effort depends only on her own characteristics. Since individuals may have different characteristics, efforts exerted by isolated players matter when identifying the key player to become isolated.

Notice that the key player policy is such that the planner only modifies the network by isolating a player. Then, all other players adapt their effort after the isolation but they are not allowed to change their links among them. Such an assumption is often justified by the fact that network relationships take more time to adjust than effort levels.[11]

3.2 An illustration

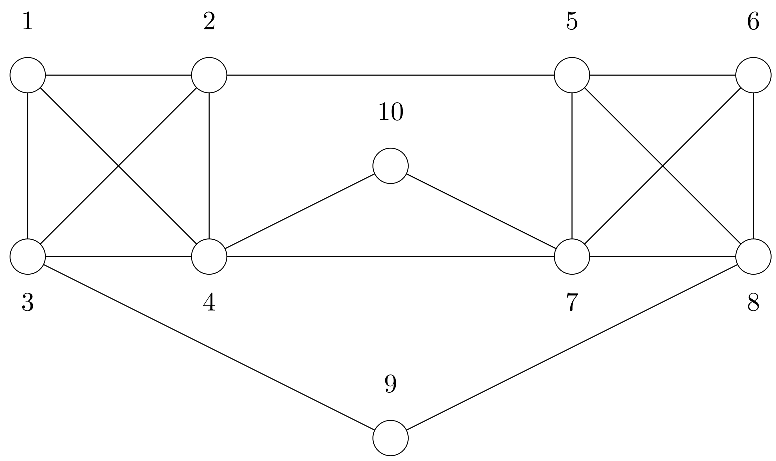


Figure 1: A network g with 10 players.

We illustrate the policy of finding and isolating the key player by means of a network g with 10 players given in Figure 1. This network structure is specific but rich enough to disentangle the effects of network spillovers, social conformity and players' characteristics on the resulting equilibrium outcomes and the key player policy. This network connects all players, and if some player is removed from the network, all other players remain connected. Let $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$, and so $\gamma = 2/3$. Players 4 and 7 are the most central players in the network, while player 9 is the least central player.[12] To understand the role played by each player's attribute and her position within the network on the key player policy, we consider the following ten cases. In all ten cases the attributes of the players sum up to 100. In addition, attributes are chosen such that all players exert a

[11]Mauleon and Vannetelbosch (2016) provide a comprehensive overview of solution concepts for solving network formation games.

[12]Katz centrality measures are 0.283 for player 1(6), 0.346 for player 2(5), 0.317 for player 3(8), 0.397 for player 4(7), 0.187 for player 9 and 0.223 for player 10.

strictly positive effort at equilibrium when the network connect all players. When a player becomes isolated, she optimally does a strictly positive effort if her attribute is positive while she exerts no effort if her attribute is negative. Players with negative attributes may be interpreted as perpetrators.

A All players have the same attribute: $y_i = 10$ for all $i \in N$.

B Players 3 and 8 have polarized attributes: $y_3 = y_8 = 2$, and $y_i = 12$ for all $i \in N \setminus \{3, 8\}$.

C Players 4 and 7 have polarized attributes: $y_4 = y_7 = 2$, and $y_i = 12$ for all $i \in N \setminus \{4, 7\}$.

D Players have decreasing attributes: $y_1 = 19$, and $y_{i+1} = y_i - 2$ for all $i \in N \setminus \{1\}$.

E Players have increasing attributes: $y_1 = 1$, and $y_{i+1} = y_i + 2$ for all $i \in N \setminus \{1\}$.

F Attributes are degree-based: $y_i = \#N_i(100/\sum_{j \in N_j})$. That is, $y_1 = y_6 = 8,33$, $y_2 = y_3 = y_5 = y_8 = 11,11$, $y_4 = y_7 = 13,89$, and $y_9 = y_{10} = 5,56$.

G Players 3 and 8 have different negative attributes: $y_3 = -1$, $y_8 = -3$, and $y_i = 13$ for all $i \in N \setminus \{3, 8\}$.

H Players 3 and 8 have the same negative attribute: $y_3 = y_8 = -2$ and $y_i = 13$ for all $i \in N \setminus \{3, 8\}$.

I Players 4 and 7 have different negative attributes: $y_4 = -1$, $y_7 = -3$, and $y_i = 13$ for all $i \in N \setminus \{4, 7\}$.

J Players 4 and 7 have the same negative attribute: $y_4 = y_7 = -2$ and $y_i = 13$ for all $i \in N \setminus \{4, 7\}$.

For the policy of finding and isolating the key player (P1), Table 1 computes, for each of the ten cases, the value of effort x_i exerted by each player,[13] the total effort $\sum_{i=1}^{10} x_i = X_{\text{P1}}$, the key player(s), and the total effort when a player is randomly isolated (X_{RP1}). For the benchmark, i.e. no policy is implemented (NP), Table 2 computes, for each of the ten cases, the value of effort x_i exerted by each player and the total effort $\sum_{i=1}^{10} x_i = X_{\text{NP}}$. The two bottom lines of Table 1 give us the relative performance of isolating the key player both with respect to doing nothing ($X_{\text{P1}}/X_{\text{NP}}$) and with respect to isolating randomly a player ($X_{\text{P1}}/X_{\text{RP1}}$).

First, we notice that the key player is not necessarily the player who is the most central within the network nor the player who is doing less effort than all other players. For instance, in the situation where players 3 and 8 have negative attributes (case H),

[13] Given some players are symmetric in the network, there could be more than one negative key player. Since the planner can only isolate a single player, we give in bold the effort done by the selected negative key player who is being isolated.

	A	B	C	D	E	F	G	H	I	J
x_1	20	23.34	23.46	28.79	14.10	21.70	20.72	21.22	20.41	25.19
x_2	20	22.95	23.01	27.50	**1**	22.20	21.64	21.99	21.41	24.52
x_3	20	**1**	23.37	26.74	16.77	22.20	17.36	17.42	21.34	25.05
x_4	**5**	23.06	**2**	24.33	18.53	22.68	22.23	23.11	17.09	**0**
x_5	20	20.89	20.72	22.11	23.35	22.20	24.53	24.79	24.61	21.08
x_6	20	20.23	20.01	19.78	23.71	21.70	25.08	25.34	25.24	20.01
x_7	20	21.31	17.64	19.24	23.61	22.68	24.69	25.23	**0**	16.46
x_8	20	17.83	20.67	**1.50**	24.75	22.20	**0**	**0**	25.11	21.01
x_9	20	16.55	22.68	23.83	22.84	**2.78**	15.57	15.28	24.15	24.02
x_{10}	20	22.79	16.43	18.19	23.71	21.60	24.31	25.11	15.40	14.64
$\sum_{i=1}^{10} x_i$	185	189.95	188.99	212.00	192.39	201.96	196.13	193.42	194.78	191.98
Key pl.	N	3,8	4,7	8	2	9	8	3,8	7	4,7
X_{RP1}	185	179.52	172	192.45	173.05	197.77	177.88	177.88	166.60	166.60
$X_{\text{P1}}/X_{\text{NP}}$	0.925	0.976	1.017	1.005	1.018	0.942	1.022	1.007	1.089	1.074
$X_{\text{P1}}/X_{\text{RP1}}$	1.000	1.058	1.099	1.102	1.112	1.021	1.103	1.087	1.169	1.152

Table 1: Efforts, total efforts, key players and relative performance for the policy of isolating the key player (P1).

player 8 (or 3) is the key player but players 4 and 7 are more central while player 9 is exerting less effort than player 8. Player 8 (or 3) turns to be the key player because (i) players 4 and 7 are doing a much higher effort than player 8 and (ii) player 9 who is the least central player is the only one to do less effort than player 8. Thus, finding the key player often deviates from simply selecting either the most central player or the player who is exerting the least effort.

Second, we observe that, once the key player is isolated, the efforts exerted by all other players increase while the effort run by the isolated player drastically decreases. Therefore, the total effort exerted by all players (including the key player) may increase or decrease when isolating the key player. When all players have the same attribute, they all exert the same effort and it is useless to isolate some player. If the attribute of a given player decreases (increases) then she will do less (more) effort except if the attributes of her neighbors substantially increase (decrease). Suppose that we start from homogeneous attributes (case A) and we substantially decrease the attributes of players 3 and 8 while we slightly increase the attributes of all other players (case B). Then, players 3 and 8 will exert much less effort while all other players will do more effort with the exception of player 9. In fact, player 9 is decreasing her effort, though her greater attribute, because she is only surrounded by players 3 and 8 whose attributes considerably decrease.

	A	B	C	D	E	F	G	H	I	J
x_1	20	19.87	19.21	28.18	11.82	21.32	20.09	19.80	19.09	18.82
x_2	20	20.21	19.61	26.69	13.31	21.86	20.52	20.32	19.60	19.42
x_3	20	17.52	19.84	25.46	14.54	21.14	16.72	16.27	19.97	19.76
x_4	20	20.67	16.00	23.64	16.36	22.38	21.18	21.01	14.31	14.00
x_5	20	20.21	19.61	19.86	20.14	21.86	20.12	20.32	19.23	19.42
x_6	20	19.87	19.21	17.51	22.49	21.32	19.51	19.80	18.55	18.82
x_7	20	20.67	16.00	17.57	22.43	22.38	20.84	21.01	13.69	14.00
x_8	20	17.52	19.84	16.37	23.63	21.14	15.82	16.27	19.55	19.76
x_9	20	16.34	21.23	18.28	21.72	19.65	14.52	14.52	21.84	21.84
x_{10}	20	21.78	15.33	17.40	22.60	21.40	22.67	22.67	13.00	13.00
$\sum_{i=1}^{10} x_i$	200	194.66	185.89	210.96	189.04	214.44	191.99	191.99	178.83	178.83

Table 2: Efforts and total efforts under the no policy (NP).

We next analyze the relative performance of the policy of finding and isolating the key player (P1) with respect to the no policy (NP). The objective of the planner is to increase the total effort, i.e. the sum of all effort levels exerted by all players (including the key player). If $X_{\text{P1}}/X_{\text{NP}}$ is strictly greater than 1 then isolating the key player is beneficial for the society. As already mentioned, when attributes are homogeneous, it is always better not to isolate any player. In addition, when attributes are positive and correlated with the centrality of the players, it is better to avoid isolating some player. For instance, when each player's attribute is simply her centrality degree (case F), the key player turns to be player 9 who is the least central player. Since player 9 is only surrounded by two very central players, she is exerting a relatively high effort. But, once she becomes isolated, her effort level drops quite substantially, and this huge decrease is not compensated by the greater effort levels done by all other players. In all other cases, implementing the policy of isolating the key player would increase the total effort exerted by all players. When players 4 and 7 have low attributes while all other players have a much higher attribute (case C), players 4 and 7 exert a low effort and are the most central players. Hence, isolating either player 4 or player 7 will increase the total effort done by all players. Indeed, by isolating such a player who is central and is exerting a low effort, the planner is able to push up the total effort because the isolated player was very influential and was influencing negatively all other players, especially her numerous neighbors. Once she is isolated, her former neighbors are now more influenced by players who are exerting higher effort levels. In general, if the most central players do low effort levels compared to other players, then they are probably the key players, and isolating them is likely to raise the total effort. When players 3 and 8 have low attributes while all other have a much higher attribute (case B), players 3 and 8 exert a low effort but they are not the most central players and so they are not influencing enough negatively the other players to lead to a substantial increase of the effort levels done by all other players once they are isolated. Hence, it is better not to isolate the key player in case B. When players have decreasing attributes (case D), the key player is player 8 rather than the most central player 7. The reason is that player 7 is a neighbor of player 4 who has a much higher attribute and is linked to players with an even higher attribute, while player 8 is not linked to player 4. In fact, player 8 exerts the lowest effort level among all players and she is sufficiently central to be optimally isolated, leading to an increase of the total effort. When two players have negative attributes (cases G, H, I, J), the key player is the one with the lowest attribute since this player is sufficiently central, and isolating this player substantially increases the total effort (up to 7%).

To assess the relevance of the key player policy (P1), we also compare the increase in total effort following the isolation of the key player with respect to what would happen if the target is selected randomly (RP1). We observe that, in all cases with the exception of homogeneous attributes, the key player policy increases the total effort by at least 2% and by at most 17% compared to the total effort done when the player to be isolated is chosen randomly. Indeed, we have that $X_{\mathrm{P1}}/X_{\mathrm{RP1}}$ is greater than one. In addition, isolating randomly a player (RP1) always decreases the total effort with respect to doing nothing (NP). Thus, one should never target and isolate randomly a player.

3.3 More about the performance of isolating the key player

From Table 1, one could be inclined to conclude that the policy of isolating the key player always increases the total effort when some players have negative attributes. In Table 3 we report the relative performance of the key player policy ($X_{\mathrm{P1}}/X_{\mathrm{NP}}$) for situations where a single player has a negative attribute. Notice that the key player is always the player with a negative attribute. We observe that the relative performance of P1 with respect to NP increases with the centrality of the player who has a negative attribute. When player 4 (or 7) who is the most central player has a negative attribute, the total effort increases by 9.3% from isolating player 4. However, when player 9 who is the least central player has a negative attribute, isolating player 9 would reduce the total effort by 2.5%. The reason is that player 9 is not influencing enough players while she is herself influenced by players who have positive attributes and are themselves mostly influenced by players with positive attributes. Thus, isolating the key player, even if she has a negative attribute, is not always beneficial for the society.

In Table 4 we analyze more deeply the relative performance of the key player policy (P1) when two players have negative attributes. We observe that, when player 4 who is the most central player has a negative attribute, she is the key player. Then, isolating player 4 always increases the total effort with the highest increase obtained when player 10 is the other player to have a negative attribute. Indeed, once player 4 is isolated, player 10 is only influenced directly by player 7 who is now the most central player and

	$y_1 = -3.5$ $y_i = 11.5$ $(i \neq 1)$	$y_2 = -3.5$ $y_i = 11.5$ $(i \neq 2)$	$y_3 = -3.5$ $y_i = 11.5$ $(i \neq 3)$	$y_4 = -3.5$ $y_i = 11.5$ $(i \neq 4)$	$y_9 = -3.5$ $y_i = 11.5$ $(i \neq 9)$	$y_{10} = -3.5$ $y_i = 11.5$ $(i \neq 10)$
X_{P1}	207	207	207	207	207	207
X_{NP}	205.06	197.86	195.99	189.41	211.16	212.18
Key player	1	2	3	4	9	10
$X_{\text{P1}}/X_{\text{NP}}$	1.009	1.046	1.056	1.093	0.980	0.976

Table 3: Relative performance of the isolating key player policy (P1) with respect to the centrality of the player with a negative attribute.

has a positive attribute. When players 2 and 9 have negative attributes, player 2 is the key player and isolating player 2 increases the total effort by 1.8%. However, if player 1 has a negative attribute instead of player 2, player 1 is not central enough to lead to an increase of the total effort when being isolated. So, again the planner does not always have incentives to isolate the key player, even if she has a negative attribute.

	$y_2 = y_4 = -2$ $y_i = 13$ $(i \neq 2, 4)$	$y_4 = y_5 = -2$ $y_i = 13$ $(i \neq 4, 5)$	$y_4 = y_9 = -2$ $y_i = 3$ $(i \neq 4, 9)$	$y_4 = y_{10} = -2$ $y_i = 13$ $(i \neq 4, 10)$
X_{P1}	202.21	196.35	212.59	220.75
X_{NP}	193.85	187.27	200.58	201.59
Key player	4	4	4	4
$X_{\text{P1}}/X_{\text{NP}}$	1.043	1.048	1.060	1.095
	$y_2 = y_9 = -2$ $y_i = 13$ $(i \neq 2, 9)$	$y_1 = y_9 = -2$ $y_i = 13$ $(i \neq 1, 9)$	$y_3 = y_5 = -2$ $y_i = 13$ $(i \neq 3, 5)$	
X_{P1}	212.80	213.69	198.12	
X_{NP}	209.02	216.23	193.85	
Key player	2	1	3	
$X_{\text{P1}}/X_{\text{NP}}$	1.018	0.988	1.022	

Table 4: Relative performance of the isolating key player policy (P1) when two players have negative attributes.

The most interesting situation arises when players 3 and 5 have negative attributes while all other players have positive attributes. In terms of centrality, player 5 is more central than player 3. However, it turns out that player 3 is the key player. Player 3 is the key player because she is linked to player 9 who is only linked to players 3 and 8, and so player 9 is strongly influenced by player 3. Isolating player 3 induces a substantial increase of the effort exerted by player 9. Player 5 is not the key player because all her neighbors are also influenced by many players with positive attributes. Thus, the most central player with a negative attribute is not necessarily the key player to be isolated. Hence, a player is more likely to become the key player to be isolated if (i) she is quite influential (i.e. she has many links), (ii) she is exerting a low effort (i.e. she has a low attribute), and (iii) her neighbors are strongly influenced by her (i.e. her neighbors have few links).

3.4 Finding the key player in presence of benevolent players

Suppose now that player 10 is a benevolent player. In Table 5 we report the relative performance of the key player policy ($X_{\text{P1S}}/X_{\text{NPS}}$) for situations where a single player has a negative attribute and player 10 is a benevolent player.[14] Comparing Table 5 with Table 3 we observe that the presence of a benevolent player (or social worker) only slightly reduces the relative performance of the isolating key player policy with respect to isolating nobody. The negative key player remains still the player with a negative attribute and the relative performance still increases with the centrality of the player who has a negative attribute, except for player 4. The reason is that, once player 10 who is a neighbor of player 4 is a benevolent player, player 4 is directly influenced by the benevolent player and so she is becoming less central and influential than without benevolent players.

	$y_1 = -3.5$ $y_i = 11.5$ $(i \neq 1)$	$y_2 = -3.5$ $y_i = 11.5$ $(i \neq 2)$	$y_3 = -3.5$ $y_i = 11.5$ $(i \neq 3)$	$y_4 = -3.5$ $y_i = 11.5$ $(i \neq 4)$	$y_9 = -3.5$ $y_i = 11.5$ $(i \neq 9)$
X_{P1S}	214.21	213.68	213.85	203.57	216.07
X_{NPS}	214.65	207.31	205.48	196.75	221.15
Key player	1	2	3	4	9
$X_{\text{P1S}}/X_{\text{NPS}}$	0.998	1.031	1.041	1.035	0.977

Table 5: Relative performance of the isolating key player policy (P1S) with respect to the centrality of the player with a negative attribute when player 10 is a benevolent player.

In Table 6 we analyze the relative performance of the key player policy (P1S) when two players have negative attributes and player 10 is a benevolent player. Comparing

[14]The effort level of the benevolent player 10 is obtained by multiplying by 1.25 her optimal effort she would exert if she was a standard player. For instance, when $y_2 = -3.5$ and $y_j = 11.5$ for all $j \neq 2$, $x^*_{10} = 21.24$ and her effort as a benevolent player is simply $x_s = 1.25 \times 21.24 = 26.56$.

	$y_2 = y_4 = -2$ $y_i = 13$ $(i \neq 2, 4)$	$y_4 = y_5 = -2$ $y_i = 13$ $(i \neq 4, 5)$	$y_4 = y_7 = -2$ $y_i = 3$ $(i \neq 4, 7)$	$y_4 = y_9 = -2$ $y_i = 13$ $(i \neq 4, 9)$
X_{P1S}	198.84	195.57	194.30	209.79
X_{NPS}	195.16	195.16	184.61	209.00
Key player	2	4	4,7	4
$X_{\text{P1S}}/X_{\text{NPS}}$	1.019	1.002	1.052	1.004
	$y_2 = y_9 = -2$ $y_i = 13$ $(i \neq 2, 9)$	$y_1 = y_9 = -2$ $y_i = 13$ $(i \neq 1, 9)$	$y_3 = y_5 = -2$ $y_i = 13$ $(i \neq 3, 5)$	$y_3 = y_8 = -2$ $y_i = 13$ $(i \neq 3, 8)$
X_{P1S}	220.94	222.26	206.05	201.17
X_{NPS}	219.55	226.91	203.90	202.07
Key player	2	1	3	3,8
$X_{\text{P1S}}/X_{\text{NPS}}$	1.006	0.980	1.011	0.996

Table 6: Relative performance of the isolating key player policy (P1S) when two players have negative attributes and player 10 is a benevolent player.

Table 6 with Table 4 we observe that, when players 2 and 4 have negative attributes, player 2 becomes now the key player to be isolated instead of player 4. Player 4 is directly influenced by the benevolent player 10 and player 2, while player 2 is only indirectly influenced by the benevolent player 10. In addition, player 2 directly influences the same number of players as player 4. Isolating player 2 means that player 4 will be even more influenced by the benevolent player 10 and no more influenced at all by player 2. Isolating player 4 would not increase the influence of the benevolent player 10 on player 2. However, when players 1 and 4 have negative attributes, player 4 remains the key player to be isolated since player 1 is still less central and influential than player 4 even though player 4 is closer to the benevolent player. Thus, in the presence of benevolent players, a player is more likely to become the key player to be isolated if she is influencing negatively many other players and she is not too influenced by benevolent players.

4 Discussion

4.1 Turning a player into a benevolent player

The planner's objective is now to find the key player, that is, the rational player who once turned into a benevolent player (or social worker) generates the highest possible increase in aggregate effort level.[15] Benevolent players choose the effort x_S while rational players still maximize their utility. Let

$$\alpha_j^{[+i]} = \frac{\phi_j}{\lambda_2+\lambda_3} + \gamma \cdot \left(\sum_{k\in S} g^*_{jk} x_S + g^*_{ji} x_S \right) = \alpha_j + \gamma \cdot g^*_{ji} \cdot x_S$$

where player i is the one who is turned into a benevolent player. Since all players keep their links, there is no change in the original network and in the contextual effects of the players.

Proposition 3. *Assume that each player has at least two links. Then, turning a player into a benevolent player increases the total effort in the game by the benevolent change inter-centrality:*

$$\begin{aligned}
\delta_i^2(g^*,\alpha) &= \underbrace{\sum_{\substack{k\in N\\ k\in N_i}} c_k(g^{*[-i]},\alpha^{[+i]}) - c_k(g^{*[-i]},\alpha)}_{\text{Benevolent player change}} - \underbrace{\frac{b_i(g^*,\alpha)}{m_{ii}} \sum_{j\in N} m_{ij}}_{\text{Intercentrality of player } i} \\
&= \sum_{\substack{k\in N\\ k\in N_i}} \sum_{j\in N} m_{jk}^{[-i]} \cdot \gamma \cdot g^*_{kl} \cdot x_S - \frac{b_i(g^*,\alpha)}{m_{ii}} \sum_{j\in N} m_{ij}.
\end{aligned}$$

Proof. Under the assumption that each player has at least two links, the effect on the total effort of turning a rational player into a benevolent player is given by

$$\begin{aligned}
&B(g^{*[-i]},\alpha^{[+i]}) - B(g^*,\alpha) = \sum_{j=1}^n \sum_{k=1}^n m_{jk}^{[-i]} \alpha_k^{[+i]} - \sum_{j=1}^n \sum_{k=1}^n m_{jk}\alpha_k \\
&= \sum_{j=1}^n \Big(\sum_{\substack{k=1\\ k\neq i}}^n m_{jk}^{[-i]} \alpha_k^{[+i]} - m_{jk}\alpha_k \Big) - \sum_{j=1}^n m_{ji}\alpha_i \\
&= \sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i\\ k\in N_i}}^n m_{jk}^{[-i]} \alpha_k^{[+i]} - m_{jk}\alpha_k + \sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i\\ k\notin N_i}}^n m_{jk}^{[-i]} \alpha_k^{[+i]} - m_{jk}\alpha_k - \sum_{j=1}^n m_{ji}\alpha_i \\
&= \sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i\\ k\in N_i}}^n m_{jk}^{[-i]} \alpha_k^{[+i]} - m_{jk}^{[-i]}\alpha_k + m_{jk}^{[-i]}\alpha_k - m_{jk}\alpha_k + \sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i\\ k\notin N_i}}^n (m_{jk}^{[-i]} - m_{jk})\alpha_k - \sum_{j=1}^n m_{ji}\alpha_i \\
&= \sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i\\ k\in N_i}}^n m_{jk}^{[-i]} (\alpha_k^{[+i]} - \alpha_k) - \Big(\sum_{j=1}^n \sum_{\substack{k=1\\ k\neq i}}^n (m_{jk} - m_{jk}^{[-i]})\alpha_k + \sum_{j=1}^n m_{ji}\alpha_i \Big)
\end{aligned}$$

[15] Davis and Davis (2007) provide a wide range of options that bystanders can use before or during or after bullying situations: e.g. specific techniques for teaching empathy and social problem solving skills, limiting the rewards of bullying behavior, and building a partnership between students and staff to create a positive and inclusive peer culture.

$$= \sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n}\sum_{j=1}^{n} m_{jk}^{[-i]}(\alpha_k^{[+i]}-\alpha_k) - \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij}$$

$$= \sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n} c_k(g^{*[-i]},\alpha^{[+i]}) - c_k(g^{*[-i]},\alpha) - \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij}$$

$$= \sum_{\substack{k=1\\k\neq i\\k\in N_i}}^{n}\sum_{j=1}^{n} m_{jk}^{(-i)}\cdot\gamma\cdot g_{kl}^*\cdot x_S - \frac{b_i(g^*,\alpha)}{m_{ii}}\sum_{j=1}^{n} m_{ij},$$

where the fourth equality follows from $\alpha_i^{[+i]} = \alpha_k$ and the sixth equality is obtained as in the Proof of Proposition 2. □

4.2 Training a player for helping victims

The planner's objective is now to find the key player, that is, the rational player who once trained generates the highest possible increase in aggregate effort level.[16] Let the training attribute be the t-th entry in $\mathbf{y}$. For player i being trained, let $\alpha_j^{i,\tau}$ be defined as

$$\alpha_j^{i,\tau} = \frac{\phi_j^{i,\tau}}{\lambda_2+\lambda_3} + \gamma\cdot\left(\sum_{k=n+1}^{n+s} g_{jk}^* x_S\right) = \alpha_j + \frac{\phi_j^{i,\tau}-\phi_j}{\lambda_2+\lambda_3},$$

with

$$\phi_i^{i,\tau} = \mathbf{y}_i^{\tau\prime}\cdot\beta_1 + \left(\sum_{j=1}^{n+s} g_{ij}^*\cdot\mathbf{y}_j'\right)\cdot\beta_2 + \xi + \epsilon_i = \phi_i + \tau\cdot\beta_1^t, \text{ and}$$

$$\phi_j^{i,\tau} = \mathbf{y}_j'\cdot\beta_1 + \left(g_{ji}^*\mathbf{y}_i^{\tau\prime} + \sum_{\substack{k=1\\k\neq i}}^{n+s} g_{jk}^*\cdot\mathbf{y}_k'\right)\cdot\beta_2 + \xi + \epsilon_j = \phi_j + g_{ij}^*\tau\cdot\beta_2^t \text{ for } j\neq i.$$

Proposition 4. *Assume that all players have at least one link. Then, training player i by τ, i.e. increasing y_i^t to $y_i^t+\tau$, increases the total effort of the game by*

$$\delta_i^3(g^*,\alpha) = \sum_{\substack{k=1\\k\in N_i\cup\{i\}}}^{n} c_k(g^*,\alpha^{i,\tau}) - c_k(g^*,\alpha) = \frac{\tau}{\lambda_2+\lambda_3}\sum_{j\in N}\left(m_{ji}\beta_1^t + \sum_{\substack{k\in N\\k\in N_i}} m_{jk}g_{ki}^*\beta_2^t\right)$$

$$= \sum_{j\in N}\Bigg(\underbrace{\frac{\tau m_{ji}\beta_1^t}{\lambda_2+\lambda_3}}_{\textit{Direct effect}} + \underbrace{\sum_{\substack{k\in N\\k\in N_i}}\frac{\tau m_{jk}g_{ki}^*\beta_2^t}{\lambda_2+\lambda_3}}_{\textit{Indirect effect}}\Bigg).$$

[16] Padgett and Notar (2013) report that peer mediation is a strategy where students themselves are taught to help resolve conflicts among their peers. In New Hampshire (USA) where middle school students have been trained as peer mediators, students who are involved in the conflict sign a contract at the end of the mediation about changing their behaviors. In addition, buddy systems encourage reporting. Students are paired with a friend or older student and these buddies would be someone on whom victims can depend for help when bullying occurs.

Proof. Under the assumption that each player has at least one link, the effect on the total effort of training player i by τ is given by

$$B(g^*, \alpha^{i,\tau}) - B(g^*, \alpha) = \sum_{j=1}^{n} \sum_{k=1}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k)$$

$$= \sum_{j=1}^{n} \Big(\sum_{\substack{k=1 \\ k \in N_i \cup \{i\}}}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k) + \sum_{\substack{k=1 \\ k \neq i \\ k \notin N_i}}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k) \Big) = \sum_{j=1}^{n} \Big(\sum_{\substack{k=1 \\ k \in N_i \cup \{i\}}}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k) + 0 \Big)$$

$$= \sum_{\substack{k=1 \\ k \in N_i \cup \{i\}}}^{n} c_k(g^*, \alpha^{i,\tau}) - c_k(g^*, \alpha) = \sum_{\substack{k=1 \\ k \in N_i \cup \{i\}}}^{n} \sum_{j=1}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k) = \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \in N_i \cup \{i\}}}^{n} m_{jk}(\alpha_k^{i,\tau} - \alpha_k)$$

$$= \sum_{j=1}^{n} \Big(m_{ji} \frac{\phi_i^{i,\tau} - \phi_i}{\lambda_2 + \lambda_3} + \sum_{\substack{k=1 \\ k \neq i \\ k \in N_i}}^{n} m_{jk} \frac{\phi_k^{i,\tau} - \phi_k}{\lambda_2 + \lambda_3} \Big) = \sum_{j=1}^{n} \Big(m_{ji} \frac{\tau \beta_1^t}{\lambda_2 + \lambda_3} + \sum_{\substack{k=1 \\ k \neq i \\ k \in N_i}}^{n} m_{jk} \frac{g_{ki}^* \tau \beta_2^t}{\lambda_2 + \lambda_3} \Big)$$

$$= \frac{\tau}{\lambda_2 + \lambda_3} \sum_{j=1}^{n} \Big(m_{ji} \beta_1^t + \sum_{\substack{k \neq i \\ k \in N_i}} m_{jk} g_{ki}^* \beta_2^t \Big) = \sum_{j \in N} \Big(\frac{\tau m_{ji} \beta_1^t}{\lambda_2 + \lambda_3} + \sum_{\substack{k \in N \\ k \in N_i}} \frac{\tau m_{jk} g_{ki}^* \beta_2^t}{\lambda_2 + \lambda_3} \Big).$$

□

The total effect of training player i can be decomposed into two parts: $\frac{\tau m_{ji} \beta_1^t}{\lambda_2 + \lambda_3}$ is the direct training effect of player i and $\sum_{\substack{k \in N \\ k \in N_i}} \frac{\tau m_{jk} g_{ki}^* \beta_2^t}{\lambda_2 + \lambda_3}$ is the indirect effect due to the contextual effects. All neighbors of player i get an increased contextual effect from player i and therefore increase their effort as well.

4.3 Comparing policies

We reconsider the network of Figure 1 and we suppose that turning a rational player into a benevolent player increases her optimal effort by 25%. For the policy of turning a player into a benevolent player (P2), Table 7 computes, for each of the ten cases, the value of effort x_i exerted by each player,[17] the total effort $\sum_{i=1}^{10} x_i = X_{\text{P2}}$, the benevolent key player(s), and the total effort when a player is randomly selected (X_{RP2}). The optimal target to be turned into a benevolent players depends on both the network and the attributes. For instance, when players 4 and 7 have low attributes compared to all other players (case C), players 4 and 7 are the most central players but they are exerting not enough effort to be turned into a benevolent player, and so it becomes optimal to target either player 3 or player 8. Players 3 and 8 are only sightly less central than players 4 and 7 but they exert a much higher effort. When players have decreasing attributes (case D), player 1 exerts the highest effort but she is not central enough to become the optimal target. Player 4 is the optimal target, and not player 7, even though they are

[17]Since there could be more than one benevolent key player and the planner can only turn a single player into a benevolent one, we give in bold the effort done by the selected player.

	A	B	C	D	E	F	G	H	I	J
x_1	21.77	21.69	20.88	30.27	12.38	23.30	21.97	21.66	20.77	20.48
x_2	21.48	21.74	20.98	28.43	13.99	23.51	22.08	21.87	20.98	20.78
x_3	21.48	19.05	**24.80**	27.22	15.08	22.81	18.29	17.83	**24.96**	**24.70**
x_4	**25.00**	**25.84**	17.19	**29.55**	17.65	**27.98**	**26.48**	**26.26**	15.51	15.18
x_5	20.60	20.83	20.03	20.57	21.80	22.53	20.75	20.95	19.66	19.84
x_6	20.50	20.38	19.50	18.10	24.47	21.87	20.04	20.32	18.84	19.10
x_7	21.15	21.86	16.39	18.93	**28.04**	23.66	22.06	22.22	14.08	14.39
x_8	20.48	18.02	20.33	16.94	25.30	21.69	16.34	16.78	20.04	20.25
x_9	20.66	17.02	23.04	19.05	22.46	20.39	15.21	15.20	23.67	23.65
x_{10}	22.05	23.90	15.86	19.83	24.90	23.70	24.85	24.83	13.53	13.52
$\sum_{i=1}^{10} x_i$	215.17	210.33	199.00	228.89	206.07	231.43	208.07	207.92	192.02	191.89
Key pl.	4,7	4,7	3,8	4	7	4,7	4	4,7	3	3,8
X_{RP2}	212.24	206.60	197.21	223.96	200.54	227.63	203.91	203.78	189.72	189.71
$X_{\text{P2}}/X_{\text{NP}}$	1.076	1.080	1.071	1.085	1.090	1.079	1.084	1.083	1.074	1.073
$X_{\text{P2}}/X_{\text{RP2}}$	1.014	1.018	1.009	1.022	1.028	1.017	1.020	1.020	1.012	1.011

Table 7: Efforts and total efforts under the policy of a benevolent key player (P2).

both the most central players, but player 4 is exerting a much higher effort. Thus, P2 does not always target neither the most central player nor the player doing the highest effort. When attribute levels are positively correlated with the centrality of players, it is optimal to target the most central player.

We observe that, the relative performance of turning a player into a benevolent player (P2) with respect to doing nothing (NP) increases the total effort by at least 7.1% and by at most 9.0%. So, it is always better for the planner to implement P2 rather than doing noting. What happens if instead of targeting the optimal player for P2, the planner chooses randomly the player to be turned into a benevolent player. We observe that P2 increases the total effort by at least 0.9% and by at most 2.8% compared to the total effort when the targeted player is chosen randomly (RP2). Thus, P2 only performs sightly better than RP2. Moreover, P2 requires the exact knowledge of the network and the attributes. Hence, if the planner cannot obtain such information or it is too costly to get it, then selecting randomly the player to become benevolent could be a good alternative. In this case, it may be preferable for the planner to select the cheaper player to be turned into a benevolent player.

We now look at the policy of training some player for helping victims (P3) in the network of Figure 1. Suppose that training a player increases her attribute by 2. Table

	A	B	C	D	E	F	G	H	I	J
x_1	20.75	20.61	19.96	28.92	12.03	22.06	20.84	20.55	19.30	19.56
x_2	20.62	20.83	20.23	27.31	13.56	22.48	21.14	20.94	19.85	20.04
x_3	20.63	18.14	20.47	26.09	14.74	21.77	17.34	16.90	20.18	20.39
x_4	**21.11**	**21.79**	**17.11**	**24.75**	16.85	**23.49**	**22.29**	**22.12**	14.80	**15.11**
x_5	20.25	20.47	19.86	20.11	20.77	22.11	20.37	20.57	19.85	19.67
x_6	20.21	20.08	19.42	17.72	23.24	21.53	19.72	20.01	19.30	19.03
x_7	20.49	21.16	16.49	18.06	**23.54**	22.86	21.33	21.50	**14.80**	14.49
x_8	20.20	17.72	20.05	16.57	24.26	21.35	16.03	16.48	20.18	19.97
x_9	20.28	16.62	21.50	18.56	22.00	19.93	14.79	14.79	22.12	22.12
x_{10}	20.87	22.65	16.20	18.27	23.46	22.27	23.54	23.54	13.87	13.87
$\sum_{i=1}^{10} x_i$	205.41	200.07	191.30	216.37	194.45	219.85	197.39	197.40	184.24	184.24
Key pl.	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7
X_{RP3}	204.00	198.66	189.89	214.96	193.04	218.44	195.99	195.99	182.83	182.83
$X_{\text{P3}}/X_{\text{NP}}$	1.027	1.028	1.029	1.026	1.029	1.025	1.028	1.028	1.030	1.030
$X_{\text{P3}}/X_{\text{RP3}}$	1.007	1.007	1.007	1.007	1.007	1.006	1.007	1.007	1.008	1.008

Table 8: Efforts and total efforts under the policy of training the key player (P3).

8 computes, for each of the ten cases, the value of effort x_i exerted by each player,[18] the total effort $\sum_{i=1}^{10} x_i = X_{\text{P3}}$, the key player(s) to be trained, and the total effort when a player is randomly trained (X_{RP3}). The optimal target to be trained for helping victims is always the most central player. We observe that, the relative performance of training a player (P3) relative to doing nothing (NP) increases the total effort by at least 2.5% and by at most 3.0%. Again, it is always better for the planner to implement P3 rather than doing nothing. What happens if instead of targeting the optimal player for P3, the planner chooses randomly the player to be trained? We observe that P3 only slightly increases the total effort by 0.7% compared to the total effort when the targeted player is chosen randomly (RP3). Thus, if collecting the information about the relationships and the attributes of the players is costly, the best choice for the planner is to select randomly the player to be trained.

The major difference between turning a player into a benevolent player and training some player for helping victims has to do with the contextual effects. Training some player increases her attribute and induces her to exert more effort. Through the contextual effects, her neighbors have also incentives to increase their effort levels. In addition, exerting higher effort levels induces neighbors to exert more effort because of the network

[18]Since there could be more than one benevolent key player and the planner can only train a single player, we give in bold the effort done by the selected player.

positive spillover effects and the social conformity. However, when some player is turned into a benevolent player, it has no impact on her attribute and the contextual effects. Only the network spillovers and the social conformity then play a role.

4.4 The cost of finding and isolating the key player

The computation of the contextual inter-centrality measure for each player relies on the knowledge of the network.[19] Thus, implementing the key player strategy obviously has its costs. The relative gains from targeting the key player instead of selecting at random increase with the variability in contextual inter-centrality measures across players. Hence, the key player strategy seems better suited for asymmetric network structures with players having quite different characteristics or attributes. Beside looking for the optimal single player to be isolated from the network to increase aggregate effort in reporting, one could easily extend the analysis to look for the optimal group to be isolated. However, implementing the negative key group strategy is much more demanding since the problem of finding a key group in a network is NP-hard (Ballester, Calvó-Armengol and Zenou, 2010). In addition of facing such computational complexity, the costs borne by the society for isolating more than one player are likely to increase with the number of players to be isolated.

4.5 Conclusion

We have studied a model where individuals are embedded in a network of relationships and they can be victims, bystanders, or perpetrators of bullying and harassment. Each individual decides non-cooperatively how much effort to exert in preventing misbehavior. Each individual's optimal effort depends on the contextual effect, the social multiplier effect and the social conformity effect. We have characterized the Nash equilibrium and we have derived an inter-centrality measure for finding the key player who once isolated increases the most the aggregated effort within the social network. An individual is more likely to be the key player if she is influencing many other individuals, she is exerting a low effort because of her characteristics, and her neighbors are strongly influenced by her. The key player policy increases substantially the aggregate effort and the targeted player should never be selected randomly. The key player without social workers is likely to remain the key player with social workers except if she is becoming less influential due to her closeness to social workers. Finally, we have considered alternative policies (e.g. training bystanders for helping victims) and we have compared them to the policy of isolating the key player.

[19] Foerster, Mauleon and Vannetelbosch (2021) study network formation games when individuals may have either public links or private ones.

Acknowledgements

We thank Emilomo Ogbe and Olivier Degomme for indepth discussions aiming at developing a model to identify network-oriented strategies for mitigating the effects of sexual and gender-based violence (SGBV) among asylum seekers. Ata Atay is a Serra Húnter Assistant Professor (Professor Lector Serra Húnter) under the Pla Serra Húnter. Ana Mauleon and Vincent Vannetelbosch are Research Director and Senior Research Associate of the National Fund for Scientific Research (FNRS), respectively. Financial support from the Belgian French speaking community ARC project 15/20-072 of Saint-Louis University - Brussels and from the Fonds de la Recherche Scientifique - FNRS research grant T.0143.18 is gratefully acknowledged. This work has been partly supported by COST Action CA16228 European Network for Game Theory.

References

[1] Ballester, C., A. Calvó–Armengol, and Y. Zenou, 2006. Who's who in networks. Wanted: The key player. *Econometrica* 74, 1403-1417.

[2] Ballester, C., A. Calvó–Armengol, and Y. Zenou, 2010. Delinquent networks. *Journal of the European Economic Association* 8, 34-61.

[3] Ballester, C., and Y. Zenou, 2014. Key player policies when contextual effects matter. *Journal of Mathematical Sociology* 38, 233-248.

[4] Boethius, S., and M. Åkerström, 2020. Revealing hidden realities: disclosing domestic abuse to informal others. *Nordic Journal of Criminology* 21, 186-202.

[5] Boguslaw, J., 2017. The key player in disruptive behavior: Whom should we target to improve the classroom learning environment, Stockholm University.

[6] Bonacich, P., 1987. Power and centrality: A family of measures. *American Journal of Sociology* 92, 1170-1182.

[7] Boucher, V., 2016. Conformism and self-selection in social networks. *Journal of Public Economics* 136, 30-44.

[8] Davis, S., and J. Davis, 2007. *From Empowering Bystanders in Bullying Prevention.* Champaign, IL: Research Press Publishers.

[9] Davis, S., and C.L. Nixon, 2013. *Youth Voice Project: Student Insights into Bullying and Peer Mistreatment.* Champaign, IL: Research Press Publishers.

[10] Debreu, G., and I.N. Herstein, 1953. Nonnegative square matrices. *Econometrica* 21, 597-607.

[11] Foerster, M., A. Mauleon, and V. Vannetelbosch, 2021. Shadow links. *Journal of Economic Theory* 197, 105325.

[12] Hawkins, D.L., D.J. Pepler, and W.M. Craig, 2001. Naturalistic observations of peer interventions in bullying. *Social Development* 10, 512-527.

[13] Jackson, M.O., and Y. Zenou, 2015. Games on networks. In: Young, P., and S. Zamir (Eds.), *Handbook of Game Theory, Vol 4*. Amsterdam: Elsevier, pp. 91-157.

[14] Landini, F., N. Montinari, P. Pin, and M. Piovesan, 2016. Friendship network in the classroom: Parents bias on peer effects. *Journal of Economic Behavior and Organization* 129, 56-73.

[15] Lee, L-F., X. Liu, E. Patacchini, and Y. Zenou, 2021. Who is the key player? A network analysis of juvenile delinquency. *Journal of Business and Economic Statistics* 39, 849-857.

[16] Lindquist, M., and Y. Zenou, 2019. Crime and networks: 10 policy lessons. *Oxford Review of Economic Policy* 35, 746-771.

[17] Mauleon, A., and V. Vannetelbosch, 2016. Network formation games. In: Bramoullé, Y., B.W. Rogers, and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*. Oxford: Oxford University Press, pp. 167-190

[18] Mazzone, A., 2020. Bystanders to bullying: An introduction to the special issue. *International Journal of Bullying Prevention* 2, 1-5.

[19] Modecki, K.L., J. Minchin, A.G. Harbaugh, N.G. Guerra, and K.C. Runions, 2014. Bullying prevelance across contexts: A meta-analysis measuring cyber and traditional bullying. *Journal of Adolescent Health* 55, 602-611.

[20] Mouttapa, M., T. Valente, P. Gallaher, I.A. Rohrbach, and B. Unger, 2004. Victimization. *Adolescence* 39, 315-355.

[21] Ogbe, E., A. Jbour, L. Rahbari, M. Unnithan, and O. Degomme, 2021. The potential role of network-oriented interventions for survivors of sexual and gender-based violence among asylum seekers in Belgium. *BMC Public Health* 21, 25.

[22] Padgett, S., and C.E. Notar, 2013. Bystanders are the Key to Stopping Bullying *Universal Journal of Educational Research* 1, 33-41.

[23] Patacchini, E., and Y. Zenou, 2012. Juvenile delinquency and conformism. The *Journal of Law, Economics, and Organization* 28, 1-31.

[24] Thornberg, T., L. Tenenbaum, K. Varjas, J. Meyers, T. Jungert, and G. Vanegas, 2012. Bystander motivation in bullying incidents: To intervene or not to intervene? *Western Journal of Emergency Medicine* 8, 247-252.

[25] Zenou, Y., 2016. Key players. In: Bramoullé, Y., B.W. Rogers, and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*. Oxford: Oxford University Press, pp. 244-274.