Key Players[*†]

Yves Zenou[‡]

November 26, 2014 (Revised February 14, 2015)

Abstract

In this chapter, we provide an overview of the literature on key players in networks. We first introduce the theoretical concept of the key player who is the agent that should be targeted by the planner so that, once removed, she will generate the highest level of reduction in total activity. We also consider another notion of key players where the planner is targeting a set of network nodes that are optimally positioned to quickly diffuse information, attitudes, behaviors or goods. Then, we examine the empirical tests of the key-player policies for criminal networks, education, R&D networks, financial networks and diffusion of microfinance. We show that implementing such a policy outperforms other standard policies such as targeting the most active agents in a network.

Key words: Key players, Katz-Bonacich centrality, crime policies, diffusion.
JEL Classification: A14, D85, K42, Z13

[*]Prepared for the *Oxford Handbook on the Economics of Networks*, Y. Bramoullé, B. Rogers and A. Galeotti (Eds.), Oxford University Press, forthcoming in 2015.

[†]I'm grateful to Sebastian Bervoets, Yann Bramoullé, Andrea Galeotti, Michael König, Mikael Priks and Brian Rogers for helpful comments. I also thank the French National Research Agency (ANR) for its support through the program ANR 13 JSH1 0009 01.

[‡]Stockholm University, IFN, and CEPR. E-mail: yves.zenou@ne.su.se.

1 Introduction

Social networks are important in several facets of our lives. For example, the decision of an agent whether or not to buy a new product, attend a meeting, commit a crime or find a job is often influenced by the choices of her friends and acquaintances.[1] Social networks have recently gained popularity with the advent of sites such as MySpace, Friendster, Orkut, Twitter, Facebook, etc. The number of users participating in these networks is large and still growing. Therefore, we need a clear understanding of the functioning of these networks. One crucial aspect of social networks **is** key players, who are those elements in the network that are considered important, in regard to some criteria. For example, identifying key players is one of the goals in an online interaction media such as blog posts (Sathik and Rasheed, 2009).

The aim of this article is to survey the research on key players in the economics of networks. First, we quickly review the key-player literature in sociology, which mostly provides centrality measures that are not microfounded. Then, we expose the economic literature by examining the theoretical models of the key player. Most of these models are within the framework of network games with strategic complementarities where the utility of the action of a player increases with the marginal increase of the actions of her neighbors or links. We expose the canonical network model of games with strategic complementarities where, at the Nash equilibrium, the action (or effort) of a player is proportional to her position in the network, as measured by her Katz-Bonacich centrality, a well-known centrality measure in sociology. For example, in the context of crime, the effort (i.e. the number of crimes) each delinquent exerts will be proportional to her position in the network, as measured by her Katz-Bonacich centrality. Then, we can determine who is the key player, which is the agent who once removed from the network generates the highest decrease in total activity. For example, in the context of crime, the key player is the criminal who, once removed, reduced total crime the most. Importantly, we show that the key player need not be the player exerting the highest effort. We are then able to extend this framework to more complex situations where, for example, competition between agents is introduced or where imperfect information on the strength of interaction is included. We also develop another framework for the key player when the focus is on the diffusion of information about a product or behaviors. In that case, the key players are the first nodes in the network that need to be targeted so that the diffusion of a product is the largest possible.

[1]For overviews of the network literature, see Goyal (2007), Jackson (2008, 2011, 2014), Ioannides (2012), Jackson, Rogers and Zenou (2014) and Zenou (2015).

More generally, which measure of centrality that is appropriate to predict behavior depends on context. In models with behavioral complementarities or where individuals may spread information, measures of an individual's centrality need to incorporate aspects of the network beyond the number of friends of an individual (i.e. degree centrality). When there are complementarities in behaviors such as in crime or education, the Katz-Bonacich centrality and the key-player intercentrality measures seem to be the right measure describing the activity of each agent. If the objective of the planner is such that the diffusion of information spreads the most, then it is not clear that we should first approach the agent with the highest Katz-Bonacich or intercentrality in a network. It may be that the agent with the highest diffusion centrality[2] would be the right node to target.

In the second part of this article, we review the empirical evidence on the key-player policies based on the theoretical models developed in the first part. A natural application is the one on criminal networks. Using both data on juvenile crime in the US and adult crime in Sweden, we show how key-player policies outperform other reasonable policies such as targeting the most prolific criminals. We then look at R&D networks where key firms are defined such that, if they exit from the market and thus from the network of collaborations, the reduction in total welfare will be the highest. Interestingly, General Motors, which was bailed out by president-elected Barack Obama in 2008, was ranked among the highest key players in the same year. We also study education and how the centrality of a group of students affects own educational outcome and financial networks where the key players are the banks that should be bailed out. Finally, we examine the diffusion of a microfinance program in India and show that targeting the individuals with the highest diffusion centrality will have the highest impact on the adoption of this microfinance program in the village.

2 Key players: Theoretical considerations

2.1 Key players: The sociological approach

The problem of identifying key players in a network is an old one, at least in the sociological literature. Indeed, one of the focuses of this literature is to propose different measures of network centralities and assert the descriptive and/or prescriptive suitability of each of these measures to different situations. Borgatti (2003, 2006) was among the first researchers to really investigate the issue of key players, which is based on explicitly measuring the contribution of a set of agents to the cohesion of a network. The basic strategy is to take

[2] A concept introduced by Banerjee et al. (2013) and defined below.

any aggregate network property, such as density or maximum flow, and derive a centrality measure by deleting nodes and measuring the change in the network property. Measures derived in this way have been called "vitality measures" (see Koschützki et al., 2005, for a review).

To be more precise, Borgatti (2003, 2006) identifies two key-player problems. First, he puts forward the "Key Player Problem/Negative" (KPP-Neg), which is defined in terms of the extent to which the network depends on its key players to maintain its cohesiveness. It is a "negative" problem because it measures the amount of *reduction* in cohesiveness of the network that would occur if the nodes were not present. Borgatti gives examples of such problems. In public health, a key-player problem is whenever the planner needs to select a subset of population members to immunize or quarantine in order to optimally contain an epidemic. In the military or criminal justice context, the key-player problem arises when a planner needs to select a small number of players in a criminal network to be neutralized (e.g., by arresting, exposing or discrediting) in order to maximally disrupt the network's ability to mount coordinated action.

The second key player problem identified by Borgatti (2003, 2006) is called the "Key Player Problem/Positive" (KPP-Pos) where the planner is looking for a set of network nodes that are optimally positioned to quickly *diffuse* information, attitudes, behaviors or goods. In a public health context, a health agency will need to select a small set of population members to use as seeds for the diffusion of practices or attitudes that promote health, such as using bleach to clean needles in a population of drug addicts. In a military or criminal justice context, the planner needs to select an efficient set of agents to surveil, to turn into something else (double-agents for example), or feed misinformation.

To identify key players, one uses centrality or the position of nodes in networks. There are numerous different ways of quantifying centrality, each having its distinct importance and logic. A basic one is simply counting the number of connections of a given node, or its degree. But there are also many richer definitions that keep track of how close a given node is to others, on average, or whether a given node is a critical connector on paths between other nodes, or whether a given node is well connected to other important nodes. Let us give a more precise definition of the most prominent centrality measures.

Degree centrality simply measures the number of links of each agent and captures a direct measure of popularity. *Betweenness centrality* of a given agent is equal to the number of shortest paths between all pairs of agents that pass through the given agent. In other words, an agent is central if she lies on several shortest paths among other pairs of agents. Betweenness centrality thus captures the importance as an intermediary. Such central agents have control over the flow of information in the network, which is related to the notion of *structural holes* developed by Burt (1992), who postulates that social capital is created by a network where people can broker connections between otherwise disconnected segments of the network. *Closeness* and *decay centrality* are measures of how close an agent is to all other agents in the network. The most central agents can quickly interact with all others because they are close to all others. These measures of centrality capture how easily an individual reaches others, i.e., how informed a given individual is in the context of information flows. *Eigenvector* centrality is a measure of the influence of an agent in a network. It assigns relative scores to all agents in the network based on the concept that connections to high-scoring agents contribute more to the score of the agent in question than equal connections to low-scoring agents. It thus captures indirect reach so that being well-connected to well-connected other agents makes you more central. Google's PageRank is a variant of the eigenvector centrality measure. Finally, *Katz-Bonacich centrality* (due to Katz, 1953 and Bonacich, 1987) takes all possible paths in a network (not only the shortest ones) but puts a lower weight on nodes that are further away from the agent.[3] As a result, Katz-Bonacich centrality captures the influence on friends and their friends. If there are strong network externalities, it can be shown that Katz-Bonacich centrality becomes proportional to eigenvector centrality (see Wasserman and Faust, 1994, Chap. 5.2).[4] Therefore, these different measures fundamentally capture different aspects of centrality.[5]

2.2 Key players: The economic approach

The standard centrality measures described above only consider network structure and do not take additional information into account. In economics, we first study the behavior of agents and their interactions and then deduce which measure of centrality that will explain these behaviors. For this purpose, we would now like to explicitly model the *behavior* of individuals since it will help us define a key-player centrality measure. We will first focus on the "Key Player Problem/Negative" defined above and will have criminal networks in mind. In that case, the aim of the planner is to identify the key player in a (connected) network, which is the individual (criminal) to be removed from the network so that total crime is minimized. What is crucial here is that when a criminal is removed from the network, the remaining

[3] For a mathematical definition of the Katz-Bonacich centrality, see (4) below.

[4] See Dequiedt and Zenou (2014) who propose an axiomatic approach to derive the degree, eigenvector and Katz-Bonacich centralities.

[5] For a review of the different existing centrality measures, see Wasserman and Faust (1994) and Jackson (2008).

criminals will optimally react by adjusting their criminal effort. As a result, we first need to define the way in which criminals exert their crime effort and examine how the key-player policy affects their criminal activities. We will then focus on "Key Player Problem/Positive" and study the role of key players in diffusion. In that case, the planner wants to target a set of network nodes that are optimally positioned to quickly *diffuse* information, attitudes, behaviors or goods.

2.2.1 Games with strategic complementarities

Although games on networks can take many forms, there are two prominent and broadly encompassing classes of games.[6] The distinction between these types of games relates to whether a given player's relative payoff to taking an action is *increasing* or *decreasing* in the set of neighbors who take this action. The first class of games on networks, of which coordination games are the canonical example, are games of *strategic complements.* In games of strategic complements, an increase in the actions of other players leads a given player's higher actions to have relatively higher payoffs as compared to that player's lower actions (Ballester et al., 2006, 2010). Games of *strategic substitutes* are such that the opposite is true: an increase in other players' actions leads to relatively lower payoffs to higher actions of a given player (Bramoullé and Kranton, 2007; Bramoullé, Kranton and D'Amours, 2014).

For criminal behaviors, it seems relatively natural to consider games with strategic complementarities since the higher are my friends' criminal efforts, the higher is my marginal utility of exerting criminal effort. Indeed, there is no formal way of learning to become a criminal, no proper "school" providing an organized transmission of the objective skills needed to undertake successful criminal activities. Given this lack of formal institutional arrangement, the most natural and efficient way to learn to become a criminal is through the interaction with other criminals. Delinquents learn from other criminals belonging to the same network how to commit crime in a more efficient way by sharing the know-how about the "technology" of crime. Another way of understanding why strategic complementarities are important in crime is that the perception that one's peers will or will not disapprove can exert a stronger influence in committing crime. Indeed, if delinquency is seen as a badge of honor in a population (Wilson and Herrnstein, 1985; Kahan, 1997; Silverman, 2004), then the delinquents who see others committing crimes infer that their peers value law-breaking. They are then more likely to break the law themselves, which leads other juveniles to draw the same inference and engage in the same behavior. In this respect, violence and crime can

[6] For a complete overview of the literature on games on networks, see Jackson and Zenou (2014) and the chapter in this handbook by Bramoullé and Kranton (2015).

become status-enhancing.

Following Calvó-Armengol and Zenou (2004) and Ballester, Calvó-Armengol and Zenou (2006, 2010), we would like to examine a simple network model with strategic complementarities in crime effort.[7] For this purpose, consider a game where $N = \{1, \ldots, n\}$ is a finite set of agents in network $\mathbf{g}$. We represent these social connections by a graph $\mathbf{g}$, where $g_{ij} = 1$ if agent i is connected to agent j and $g_{ij} = 0$ otherwise. Links are taken to be reciprocal, so that $g_{ij} = g_{ji}$.[8] By convention, $g_{ii} = 0$. We denote by $\mathbf{G}$ the $n \times n$ adjacency matrix with entry g_{ij}, which keeps track of all direct connections. In criminal activities, agents i and j share their knowledge about delinquent activities if and only if $g_{ij} = 1$. Each agent i decides how much effort to exert on crime, denoted by $y_i \in \mathbb{R}_+$. The utility of each agent i providing effort y_i in network $\mathbf{g}$ is given by:

$$u_i\left(\mathbf{y}, \mathbf{g}\right) = \alpha_i y_i - \frac{1}{2} y_i^2 + \phi_1 \sum_{j=1}^{n} g_{ij} y_i y_j \tag{1}$$

where $\phi_1 > 0$ is the intensity of interactions and $\mathbf{y}$ is an n-dimensional vector of crime efforts. This utility has two parts. An individual part, $\alpha_i y_i - \frac{1}{2} y_i^2$ where the marginal benefits of providing criminal effort y_i are given by $\alpha_i y_i$ and increasing in own effort y_i, where α_i denotes the *exogenous heterogeneity* of agent i that captures the *observable* characteristics of individual i (e.g. sex, race, age, parental education). The second part of the utility function, $\phi_1 \sum_{j=1}^{n} g_{ij} y_i y_j$, corresponds to the *local-aggregate* effect of peers since each agent i is affected by the sum of efforts of the agents for which she has a direct connection. The higher are these number of active connections, the higher is the marginal utility of providing her own effort. This is a game with strategic complementarities since

$$\frac{\partial^2 u_i\left(\mathbf{y}, \mathbf{g}\right)}{\partial y_i \partial y_j} = \phi_1 g_{ij} \geq 0.$$

In equilibrium, each agent maximizes her utility (1) and the best-reply function, for each $i = 1, ..., n$, is given by:

$$y_i = \alpha_i + \phi_1 \sum_{j=1}^{n} g_{ij} y_j. \tag{2}$$

Denote by $\mu_1(\mathbf{G})$ the largest eigenvalue of network $\mathbf{g}$ and by $\boldsymbol{\alpha}$ the non-negative n-dimensional vector corresponding to α_i. Denote also by $\mathbf{I}_n$ the $(n \times n)$ identity matrix, $\mathbf{1}_n$ the n-

[7] The same model or an extension of it can be used for education, R&D, trade, etc. See below.

[8] This is only for the sake of the exposition. All the results go through with a directed and weighted network.

dimensional vector of ones and $\mathbf{M}(\mathbf{g}, \phi_1) \equiv (\mathbf{I}_n - \phi_1 \mathbf{G})^{-1}$. We have the following result:[9]

Proposition 1 *If $\phi_1 \mu_1(\mathbf{G}) < 1$, the network game with payoffs (1) has a unique Nash equilibrium in pure strategies given by:*

$$\mathbf{y}^* \equiv \mathbf{y}^*(\mathbf{g}) = \mathbf{b}_{\boldsymbol{\alpha}}(\mathbf{g}, \phi_1), \tag{3}$$

where $\mathbf{b}_{\boldsymbol{\alpha}}(\mathbf{g}, \phi_1)$ is the weighted Katz-Bonacich centrality defined as:

$$\mathbf{b}_{\boldsymbol{\alpha}}(\mathbf{g}, \phi_1) = (\mathbf{I}_n - \phi_1 \mathbf{G})^{-1} \boldsymbol{\alpha} = \mathbf{M}(\mathbf{g}, \phi_1)\boldsymbol{\alpha} = \sum_{k=0}^{\infty} \phi_1^k \mathbf{G}^k \boldsymbol{\alpha}. \tag{4}$$

When there is no ex ante heterogeneity, i.e. $\boldsymbol{\alpha} = \mathbf{1}$, the Katz-Bonacich centrality of agent i counts the total number of paths (not just the shortest paths) in $\mathbf{g}$ starting from i, weighted by a decay factor that decreases with the length of these paths. This is captured by the fact that the matrix $\mathbf{G}^k$ keeps track of the indirect connections in the network, i.e. $g_{ij}^{[k]} \geq 0$ measures the number of paths of length $k \geq 1$ in $\mathbf{g}$ from i to j. When there is individual heterogeneity, i.e. $\boldsymbol{\alpha} \neq \mathbf{1}$, then paths have different weights depending on where they arrive. In particular, each path is also weighted by the sum of the αs to which it corresponds. Proposition 1 shows that more central agents in the network will exert more effort. This is intuitively related to the equilibrium behavior, as the paths capture all possible feedbacks. In our case, the decay factor depends on how others' effort enters into the payoff of own effort. It is then straightforward to show that, for each individual i, the equilibrium utility is:

$$u_i(\mathbf{y}^*, \mathbf{g}) = \frac{1}{2} \left[b_{\alpha_i}(\mathbf{g}, \phi_1)\right]^2 \tag{5}$$

so that the equilibrium utility of each criminal is proportional to the square of her Katz-Bonacich centrality.

2.2.2 The key-player policy

We here focus on the "Key Player Problem/Negative" (KPP-Neg) where the key-player policy aims at removing the player (criminal) who reduces total activity (crime) in a network the most. The removal of the key player can have large effects on crime because of feedback effects or "social multipliers" (see, in particular, Kleiman, 2009; Glaeser, Sacerdote and Scheinkman, 1996; Verdier and Zenou, 2004). That is, as the fraction of individuals

[9] Throughout this paper, bold-face, lower-case letters refer to vectors while bold-face, capital letters refer to matrices.

participating in a criminal behavior increases, the impact on others is multiplied through social networks. Thus, criminal behaviors can be magnified, and interventions can become more effective.[10]

The benchmark key-player policy Formally, consider the previous model and denote by $Y^*(\mathbf{g}) = \sum_{i=1}^{n} y_i^*$ the total equilibrium level of crime in network $\mathbf{g}$, where y_i^* is the Nash equilibrium effort given by (3). Denote also by $\mathbf{g}^{[-i]}$ the network $\mathbf{g}$ without individual i. Then, in order to determine the key player, the planner will solve the following problem:

$$\max\{Y^*(\mathbf{g}) - Y^*(\mathbf{g}^{[-i]}) \mid i = 1, ..., n\}.$$

When the original delinquency network $\mathbf{g}$ is fixed, this is equivalent to:

$$\min\{Y^*(\mathbf{g}^{[-i]}) \mid i = 1, ..., n\}. \tag{6}$$

Definition 1 *Assume that $\phi\mu_1(\mathbf{g}) < 1$. The intercentrality or key-player centrality measure $d_i(\mathbf{g}, \phi_1)$ is defined as follows:*

$$d_i(\mathbf{g}, \phi_1) = \frac{b_{\alpha_i}(\mathbf{g}, \phi_1) b_{1_i}(\mathbf{g}, \phi_1)}{m_{ii}} \tag{7}$$

Ballester et al. (2006, 2010) have showed the following result:

Proposition 2 *A player i^* is the key player that solves (6) if and only if i^* is a delinquent with the highest intercentrality in $\mathbf{g}$, that is, $d_{i^*}(\mathbf{g}, \phi_1) \geq d_i(\mathbf{g}, \phi_1)$, for all $i = 1, ..., n$.*

The intercentrality measure (7) of delinquent i is the sum of i's centrality measures in $\mathbf{g}$, and i's contribution to the centrality measure of every other delinquent $j \neq i$ also in $\mathbf{g}$. It accounts both for one's exposure to the rest of the group and for one's contribution

[10] Contrary to the key player approach developed in this section, which is a *non-cooperative game theoretical approach*, Lindelauf et al. (2013) have proposed to tackle the key-player problem in *cooperative game theoretical approach*, which is particularly well suited for organized crime and terrorist organizations. In particular, Lindelauf et al. (2013) use the Shapley value as a measure of importance in cooperative games that are specifically designed to reflect the context of the terrorist organization at hand. The advantage of this approach is that both the structure of the terrorist network, which usually reflects a communication and interaction structure, and non-network features, i.e., individual based parameters such as financial means or bomb building skills, can be taken into account. See also Husslage et al. (2014). Observe that, as in the sociological approach, there is no microfoundation for the use of their proposed centrality measures.

to every other exposure. This means that the key player i^* in network $\mathbf{g}$ is given by $i^* = \arg\max_i d_i(\mathbf{g}, \phi_1)$, where

$$d_i(\mathbf{g}, \phi_1) = Y^*(\mathbf{g}) - Y^*\left(\mathbf{g}^{[-i]}\right). \tag{8}$$

Let us now discuss different extensions of this benchmark model and their implication for the key-player policy.[11]

Imperfect information Consider the same utility function as in (1) but assume that the players do not know the exact value of the synergy parameter ϕ_1 (the state of the world), which is *common* to all agents but only *partially known* by the agents. Each agent i receives a private (independent) signal s_i about the state of the world, which allows her to update her beliefs about ϕ_1. There are M different states of the world so that ϕ_1 can take M different values: $\phi_1 \in \{\phi_{11}, \ldots, \phi_{1M}\}$. There are T different signals so that agents can be of T different types, which we denote by $\tau = 1, ..., T$. Individual i, who receives signal $s_i = \tau$, computes the following conditional expected utility:

$$\begin{aligned}
\mathbb{E}\left[u_i\left(\mathbf{y}, \mathbf{g}\right) \mid \{s_i = \tau\}\right] &= \alpha_i \mathbb{E}\left[y_i \mid \{s_i = \tau\}\right] - \frac{1}{2}\mathbb{E}\left[y_i^2 \mid \{s_i = \tau\}\right] + \sum_{j=1}^{n} g_{ij} y_i \mathbb{E}\left[\phi_1 y_i y_j \mid \{s_i = \tau\}\right] \\
&= \alpha_i\, y_i\left(\tau\right) - \frac{1}{2} y_i\left(\tau\right)^2 + \sum_{j=1}^{n} g_{ij} y_i\left(\tau\right) \mathbb{E}\left[\phi_1 y_j \mid \{s_i = \tau\}\right]
\end{aligned}$$

where $\mathbb{E}[.]$ is the expectation operator. De Marti and Zenou (2014) show the existence and uniqueness of the Nash equilibrium in effort and characterize it as a function of a weighted combination of the Katz-Bonacich centrality of the players and the information matrix, which keeps tracks of all the information received by the agent about the states of the world. The authors then analyze the key-player policy where the planner plays first. They assume that the planner has a prior (which is unknown to the agents) that may be different to the one shared by the agents because the authority may have superior information. The planner needs to solve the key player problem, that is $\max\{\mathbb{E}\left[Y^*(\mathbf{g})\right] - \mathbb{E}\left[Y^*(\mathbf{g}^{[-i]}\right]) \mid i = 1, ..., n\}$, which is the difference in aggregate activity according to her prior. They show that the formula of the key player is different to (7) and that it is, in fact, a convex combination between different intercentrality measures where the weights are the different parameters related to the information structure, in particular, the priors of the agents and the planner and the posteriors of the agents.

[11] Sommer (2014) further develops the notion of key player defined in (7). He states that the formula (7) implies that when someone is removed from the network, she does not participate in any criminal activity at all. Sommer (2014) proposes another policy where the removed criminal can still exert crime but without influencing anybody in the network. She is just isolated from all other players of the game.

Network formation In the key-player formula (7), it is assumed that when someone is removed, the remaining players in the network adjust their optimal criminal efforts but cannot form new links. This makes sense in the short run but it is possible that, in the long run, there is a "rewiring" of the network so that, following the removal of a criminal, the remaining players form new links. Liu et al. (2012) propose a simple dynamic formation model that incorporates this effect. More precisely, each period of time is divided into two subperiods. In the first period, a player is randomly selected from the network and must decide *myopically* whether she wants to form a link and with whom. In the second period, all players connected in the network play the effort game where their equilibrium utility is given by (5). What is interesting here is that when a person wants to form a link, she anticipates the rewiring of the network and thus, the new equilibrium efforts or, equivalently, the new Katz-Bonacich centralities of all individuals in the network (see (3)). The chosen person will then form a link with the individual that increases her utility the most. If there is no cost of forming new links and if there is some noise in the link-formation process, König, Tessone and Zenou (2014) show that the steady-state equilibrium network will be a *nested split graph*, a well-known network in the graph-theory literature (Mahadev and Peled, 1995), which also starts to be acknowledged in the economics literature (Belhaj, Bervoets and Deroïan, 2014; Hiller, 2014; Lagerås and Seim, 2014).[12]

Liu et al. (2012) develop this dynamic model but impose a cost of man-eating links, which is specific to each individual. Thus, the steady-state equilibrium network is no longer a nested-split graph. However, the Markov process for which the state variable is the network itself is still well defined, even though it is not ergodic. In this framework, an equilibrium is reached when there is an absorbing state. The planner can then decide who the key player is by first removing an individual in the network and then letting the other individuals play the dynamic network-formation game described above. The person who reduces total crime the most in equilibrium is the key player. The latter is not necessarily the key player in the static model defined in (7).

Congestion and competition effects Another interesting extension is to generalize the utility function (1) to include competition effects so that:

$$u_i(\mathbf{y}, \mathbf{g}) = \alpha_i y_i - \frac{1}{2} y_i^2 + \phi_1 \sum_{j=1}^{n} g_{ij} y_i y_j - \rho \sum_{j=1}^{n} c_{ij} y_i y_j \tag{9}$$

[12] A graph is nested split if agents can be ordered so that $g_{kl} = 1 \Rightarrow g_{ij} = 1$ whenever $i \leq k$ and $j \leq l$. This means, in particular, that if agent i has a lower degree than agent j, then she has less neighbors in the sense of set inclusion.

where ρ is the degree of competition between agents and c_{ij} is the ijth cell of the matrix $\mathbf{C}$, which keeps track of who is in competition with whom. For example, in the case of delinquent networks, criminals benefit from other criminals who are linked to them (strategic complementarities) but are also in competition with all criminals in the neighborhood where they operate, which leads to congestion effects. Thus, the utility of i will be lower the more competition there is in the same neighborhood since there will be less to steal. In that case, the matrix $\mathbf{C}$ will keep track of who resides or operates in the same neighborhood with whom. For example, consider a star network where individual 1 is the star. Assume that there are two neighborhoods, $\mathcal{N}_1$ and $\mathcal{N}_2$, and that criminals 1 and 2 commit their crime in the same neighborhood $\mathcal{N}_1$ (for example, they sell drugs) while criminal 3 operates alone in neighborhood $\mathcal{N}_2$. Then, the adjacency matrix $\mathbf{G}$ and the competition matrix $\mathbf{C}$ can be written as:

$$\mathbf{G} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{10}$$

This model with utility (9) was originally proposed by Ballester, Calvó-Armengol and Zenou (2006) and Calvó-Armengol, Patacchini and Zenou (2009) and generalized by König, Liu and Zenou (2014) for any matrix $\mathbf{G}$ and $\mathbf{C}$. Let $\overline{\alpha} = \max\{\alpha_i \mid i \in N\}$ and $\underline{\alpha} = \min\{\alpha_i \mid i \in N\}$, with $\overline{\alpha} > \underline{\alpha} > 0$. If $\phi_1 \mu_1 (\mathbf{G}) + n \left(\frac{\rho}{1-\rho}\right)\left(\frac{\overline{\alpha}}{\underline{\alpha}} - 1\right) < 1$, then there is a unique interior Nash equilibrium defined by:

$$y_i^* = \alpha_i + \phi_1 \sum_{j=1}^{n} g_{ij} y_j^* - \rho \sum_{j=1}^{n} c_{ij} y_i^* y_j^*. \tag{11}$$

Then, we can define the key player as in (7) but where $b_{\alpha_i}(\mathbf{g}, \phi_1)$, $b_{1_i}(\mathbf{g}, \phi_1)$ and m_{ii} are now defined with respect to both $\mathbf{G}$ and $\mathbf{C}$ and *not* only with respect to $\mathbf{G}$ as in (7).

Occupational choice before effort decisions So far, we have assumed that the delinquency network was given. In some cases, though, delinquents may have opportunities outside the delinquency network. For instance, petty delinquents may consider entering the labor market and giving up delinquent activities. Here, we expand the model and endogenize the delinquency network by allowing delinquents to take a binary decision on whether to stay in the delinquency network or to drop out of it. Ballester, Calvó-Amengol and Zenou (2010) show that there always exists a subgame-perfect Nash equilibrium of this two-stage game but it is usually not unique. Then, they study the key player policy. Initially, the planner must choose a player to remove from the network. Then, players play the two-stage delinquency game. In this context, there is an added difficulty to the planner's decision. The removal of a player from the network affects the rest of the players' decisions to become active delinquents. This fact should be taken into account by the planner in order to attain an equilibrium with minimum total delinquency. Denote the fixed wage earned in the labor market by ω. Then, in order to reduce delinquency, the planner should consider two policies: the one that provides a higher ω and the key player policy. These policies are *complementary* from the point of view of their effects on total delinquency, although we are aware that they may be substitutes if we impose a budget-restricted planner who had to implement costly policies.[13]

Local average versus local aggregate network model The utility in (1) corresponds to the *local-aggregate* model because what matters for peer effects is the sum of the activity of direct friends. There is another model introduced by Patacchini and Zenou (2012) where the utility is equal to:

$$u_i(\mathbf{y}, \mathbf{g}) = a_i y_i - \frac{1}{2} y_i^2 - \frac{1}{2}\lambda \left(y_i - \sum_{j=1}^{n} g_{ij}^* y_j \right)^2 \tag{12}$$

where $g_{ij}^* = g_{ij}/g_i$. This model is called the *local-average model* because what matters for each individual i is the difference between her effort y_i and the average effort of her neighbors $\sum_{j=1}^{n} g_{ij}^* y_j$. Each individual i wants to minimize the social distance between herself and her reference group, where λ is the parameter describing the taste for conformity. Indeed, the individual loses utility $\frac{1}{2}\lambda \left(y_i - \sum_{j=1}^{n} g_{ij}^* y_j \right)^2$ from failing to conform to others. It is easily shown that there exists a unique interior Nash equilibrium given by:

$$y_i^* = \alpha_i + \phi_{12} \sum_{j=1}^{n} g_{ij}^* y_j \tag{13}$$

where $\alpha_i = a_i/(1+\lambda)$ and $\phi_{12} = \lambda/(1+\lambda)$. Interestingly, when one considers the key-player policy in the local-average model, then one can see that, in general, when the agents are *ex ante* homogeneous, which agent to remove from the network does not matter in terms of the aggregate effort level reduction, unless the agent holds a very special position in the network such that removing this agent generates isolated nodes in the network (Liu et al. 2014). If, on the contrary, the agents are ex ante heterogenous in terms of α_i, the key-player problem for

[13] See also Chen, Zenou and Zhou (2015) for a model where agents provide two activities (for example, crime and education) and where the key player is determined. Chen, Zenou and Zhou (2015) show that, by not taking into account the two activities, the planner may target the wrong key player.

the local-average network game and the general network game with utility (12) does not have an analytical solution. The difficulty is due to the fact that the row-normalized adjacency matrix of network $\mathbf{g}^{[-i]}$ is, in general, not a submatrix of the row-normalized adjacency matrix of network $\mathbf{g}$ (since $\mathbf{G}^{[-i]*}$ is not a submatrix of $\mathbf{G}^{*}$). Yet one can still determine the key player numerically using its definition given by (8), if the unknown parameters can be estimated in the best-response function (13).

Liu, Patacchini and Zenou (2014) and Liu et al. (2014) extend this model to incorporate both local-average and local-agreggate effects in the utility function. This is defined as the *hybrid model* and is given by:

$$u_i(\mathbf{y}, \mathbf{g}) = a_i y_i - \frac{1}{2} y_i^2 + \lambda_1 \sum_{j=1}^{n} g_{ij} y_i y_j - \frac{1}{2} \lambda_2 \left(y_i - \sum_{j=1}^{n} g_{ij}^{*} y_j \right)^2 . \tag{14}$$

The best-reply function for each agent i is then given by:

$$y_i^* = \alpha_i + \phi_1 \sum_{j=1}^{n} g_{ij} y_j + \phi_2 \sum_{j=1}^{n} g_{ij}^{*} y_j \tag{15}$$

where $\alpha_i = a_i / (1 + \lambda_2)$, $\phi_1 = \lambda_1 / (1 + \lambda_2)$ and $\phi_2 = \lambda_2 / (1 + \lambda_2)$. Let $g^{\max} = \max_i g_i$ denote the highest degree in network $\mathbf{g}$. Then, if $g^{\max} \phi_1 + \phi_2 < 1$, the network game with utility (14) has a unique Nash equilibrium in pure strategies given by (15). We can once more calculate the key player in this extended model by using the formula (8).[14]

2.2.3 Key players: Maximizing diffusion

The model of Section 2.2 states that when there are *strategic complementarities* in efforts between linked individuals as, for example, in crime, R&D and financial activities, then it makes sense to use the methodology of Section 2.2 to identify the key players. However, in other activities, when strategic complementarities are less important, then other methodologies could be applied. This is particularly true in a diffusion process where the planner

[14] Other extensions of the key-player policy have been considered. First, researchers have also defined "group" players so that the planner can remove more than one key player at a time (see, in particular, Ortiz-Arroyo, 2010, and Ballester, Calvó-Armengol and Zenou, 2010). Ballester, Calvó-Armengol and Zenou (2010) show that the key group problem is NP-hard from the combinatorial perspective. However, they show that the error of approximation of using a greedy algorithm (where, in each step, the key-player formula (7) is used to remove a player) compared to directly solving the key group problem is at most 36.79%. Second, König, Liu and Zenou (2014) propose a slightly different definition of the key player, which is now the agent who, once removed, reduces *total welfare* the most (and not *total activity* as in (7)). Criminal networks are a natural application but R&D or financial networks could also be another application of the key-player policy.

wants to target specific individuals (key players) to diffuse a technology or specific programs or to eradicate a virus or a sickness. In Section 2.1, this was referred to as the "Key Player Problem/Positive" (KPP-Pos) where the planner is looking for a set of network nodes that are optimally positioned to quickly *diffuse* information, attitudes, behaviors or goods.

The canonical models of diffusion (the Susceptible-Infected-Recovered or SIR model and the Susceptible-Infected-Susceptible or SIS model) constitute a good starting point for this type of issue. The diffusion process works as follows. For example, in the SIS model, individuals exist in one of two states, S (Susceptible to be infected) and I (Infected), and transition between these states is analyzed over time. Transitions from I to S occur at a given recovery rate, but transitions from S to I occur at an endogenous rate that, in particular, depends on the expected number of individuals in state I that will be met in a given period. The SIS model and its variants have been proposed as useful in understanding a wide variety of processes including such diverse applications as behaviors, information diffusion, learning dynamics, myopic best response and imitation dynamics.[15]

Galeotti and Rogers (2013) extend the SIS model to analyze the spread of a harmful state (for example, human infection of various communicable diseases, that spread through social contacts or tobacco use or an electronic virus on a computer network) through a population divided into two groups, labeled A and B. Individuals in the two groups are identical in every respect but the label. As in the SIS model, in each period, an individual interacts with k other individuals. A proportion of the interactions is with individuals from the same group while the remaining interactions are with individuals from the other group.

Their first set of results concerns the problem of a central planner attempting to eradicate the harmful state (infected state). The planner might be a governmental agency deciding how many vaccines to produce for a communicable disease and how to allocate them across the population. Alternatively, the social planner could be a governmental agency that aims at eliminating smoking in schools via an educational program and it has to decide how large the program should be and which students it should target or require to participate. How should the planner's policy depend on the structure of interactions among potential smokers? What information does the government need in order to determine the optimal size of the program (level of immunity) and how to efficiently distribute the resources to the population?

Galeotti and Rogers (2013) show that a central planner who aims for eradication optimally either divides the resources equally across groups, or concentrates entirely on one

[15] For an overview on diffusion and networks, see Jackson and Yariv (2011) and the chapter by Lamberson (2015) in this handbook.

group, depending on whether there is positive or negative assortativity, respectively. To be more precise, where there are *positive assortative interactions*, the optimal allocation involves splitting the resources equally between the two groups and the minimum budget needed to eradicate the harmful state is independent of the exact degree of positive assortativity. In contrast, when there are *negative assortative interactions*, the optimal allocation is to first exclusively focus resources on one group until it is fully immunized and then immunize part of the other group, if required. Furthermore, under negative assortative interactions, the minimum budget necessary to achieve eradication is decreasing in the intensity of assortativity, reaching its lowest level when the network of interactions is bipartite.

This analysis, though interesting, focuses more on groups and less on targeting individuals. A recurring theme in the popular discussion as well as the academic literature on social networks has been the simple idea that it would be better to focus efforts—send information, coupons, or free samples—on individuals who are influential. Galeotti and Goyal (2009) identify conditions for the content of interaction under which optimal strategy targets more or less connected individuals. In particular, they find that, in the word of mouth application, it is optimal to target individuals who get information from a few others (marginalized consumers). By contrast, in the proportional adoption externalities application, it is optimal to seed the most connected individuals (as they are unlikely to adopt via social influence). Thus, the optimality of targeting highly connected nodes very much depends on the content of social interaction.

Another recent paper by Banerjee et al. (2013) addresses the issue of diffusion of a micro-finance program in India. Their key question is: How do the network positions of the first individuals in a village to receive information about a new product affect its eventual diffusion? To answer this question, Banerjee et al. (2013) develop a model of information diffusion through a social network that discriminates between *information passing* (individuals must be aware of the product before they can adopt it, and they can learn from their friends) and *endorsement* (the decisions of informed individuals to adopt the product might be influenced by their friends' decisions). To be more precise, at each point in time t, a node (household) i can be in two possible *information* states: either she is informed so that $s_{it} = 1$ or she is uninformed so that $s_{it} = 0$. Moreover, she can be in two possible *participation* states: $m_{it} = 1$ if he/she participates and $m_{it} = 0$ if not. Note that, by definition, if $m_{it} = 1$ then $s_{it} = 1$, as one cannot participate without being informed. Let I_t be the set of newly informed nodes at time t, i.e. $I_t = \{s_{it} = 1, s_{it-1} = 0\}$. Define I^t to be the historical stock. The authors develop an algorithm that works as follows. At the beginning of the period ($t = 0$), the initial set of nodes (i.e. the *leaders*) is informed so that $s_{i0} = 1$, $\forall i \in L$

and $s_{i0} = 0$ if $i \notin L$, where $I_0 = \{i \in N : i \text{ is a leader}\}$. This is the first stage where an initial set of households is informed (*injection points*). Then, those newly informed agents decide whether or not to *participate* based on their characteristics and the participation decisions of their neighbors. To be more precise, let p_{it} denote the probability that an individual who was just informed about microfinance decides to participate, where p_{it} is a function of the individual's characteristics X_i (which can account for homophily based on observables) and peer decisions. The authors assume a logistic function so that:

$$\log\left(\frac{p_{it}}{1-p_{it}}\right) = X_i'\beta + \lambda F_{it} \tag{16}$$

where F_{it} is a fraction whose denominator is the number of i's neighbors who informed i about the program and whose numerator is the number of these individuals who participate in microfinance. If we set λ to 0, then we have a pure information model without any endorsement effects, for each $i \in I_0$. In the case of pure endorsement effects, for the initial period, we can fix $F_{i0} = 0$. Next, each $i \in I^0$ transmits to $j \in N_i$ with probability $m_{i1}\phi_1^P + (1-m_{i1})\,\phi_1^N$ so that ϕ_1^P and ϕ_1^N are, respectively, the probability that households that have been informed in previous periods pass information to each of their neighbors, independently, if they are participants (P) and if they are not (N). This is independent across i and j. Let I_1 be the set of j's neighbors who are informed via this process who were not members of I^0, and let $I(j)$ be the set of i's who informed j. Then, we iterate this process at time t so that newly informed agents are now I_t and have to decide whether to participate using the same rule as in (16) but for a different set of neighbors F_{it}. The process stops after T periods of information passing. If $\phi_1^N = 0$, so that only participating households pass information, and $T = +\infty$, this is a variant of the standard Susceptible, Infectious, Recovered (SIR) model described above. By allowing it to only operate for T periods, the authors can study what happens in finite time (since after enough rounds, everyone would be informed). They define that the key player(s) are the agents who diffuse information the best. As we will see below, they propose some centrality measures to define the key players in this context.

Finally, there is an interesting literature on *viral marketing*, which looks at how firms can use network information to better sell their products. Viral marketing may take the form of video clips, interactive Flash games, advergames, ebooks, brandable software, images, text messages, email messages, or web pages. The most commonly utilized transmission vehicles for viral messages include: pass-along based, incentive based, trendy based, and undercover based. The ultimate goal of marketers interested in creating successful viral marketing programs is to create viral messages that appeal to individuals with high *social networking potential* and that have a high probability of being presented and spread by these individuals and their competitors in their communications with others in a short period of time. In other words, one key aspect of viral marketing is to find the key consumers who have the most influence on other consumers. Indeed, in viral marketing, a company tries to use word-of-mouth effects to market a product with a limited advertising budget, relying on the fact that early adopters may convince friends and colleagues to use the product, thus creating a large wave of adoptions. While word-of-mouth effects have a history in the area of marketing that long predates the Internet, viral marketing has become a particularly powerful force in on-line domains, given the ease at which information spreads, and the rich data on customer behavior that can be used to facilitate the process.[16]

To illustrate the key-player idea in terms of the viral marketing framework, let us consider the algorithm problem posed by Domingos and Richardson (2001) to identify the "influential" sets of nodes in a network. Suppose that a firm is trying to market a new product, and wants to take advantage of word-of-mouth effects. One strategy would be as follows: this firm can collect data on the social network interactions among potential customers, chooses a set S of initial adopters, and markets the product directly to them. Then, assuming that they adopt the product, the firm will rely on their influence to generate a large cascade of adoptions, without having to rely on any further direct promotion of the product. Here, we focus on the following algorithmic problem: how does the firm choose the set S? There is a natural *influence function* $f(.)$ defined as follows: for a set S of nodes, $f(S)$ is the expected number of active nodes at the end of the process, assuming that S is the set of nodes that are initially active. From the marketer's point of view, $f(S)$ is the expected number of total sales if they get S to be the set of initial adopters. Now, given a budget k, how large can we make $f(S)$ if we are allowed to choose a set S of k initial adopters? In other words, we wish to maximize $f(S)$ over all sets S of size k. This turns out to be a hard computational problem since it is $NP-$hard to find the optimal set S. Kleinberg (2008) proposes to identify broad sub-classes of the models that are not susceptible to strong inapproximability results, and for which good approximation results can be obtained. He assumes that $f(S)$ is a submodular function, which is a type of diminishing returns property: the benefit of adding elements decreases as the set to which they are being added grows.[17]

While the previous approach focused on *influence maximization*, Hartline, Mirrokni and Sundararajan (2008) study *revenue maximization*. In their model, a buyer's decision to buy an item is influenced by the set of other buyers that own the item and the price at which the

[16] For an overview of the literature on viral marketing, see Kleinberg (2007) and the chapters by Bloch (2015) and Mayzlin (2015) in this handbook.

[17] For more details, see Kempe, Kleinberg and Tardos (2003, 2005).

item is offered. They identify a family of strategies called influence-and-exploit strategies that is based on the following idea: Initially influence the population by giving the item for free to a carefully chosen set of buyers. Then, extract revenue from the remaining buyers using a 'greedy' pricing strategy.[18]

3 Key players: Empirical results

There are very few empirical papers on key players in networks and most of them test the key-player model developed in Section 2.2.2. We will also give some empirical results on the empirical test of the model developed in Section 2.2.3.

3.1 Key player: Econometric methodology

There are two steps to empirically test the key-player policy of Section 2.2.2. First, one needs to have a credible estimate of the social multiplier ϕ_1 by estimating equation (2). Then, one must calculate the formula of the key player given in (7), using the estimated values of equation (2). Liu et al. (2012) were the first to propose an econometric methodology to test this key-player policy. In the real-world, there is often more than one (connected) network. Denote by $r = 1, ..., \bar{r}$, the network where $\bar{r}$ is the total number of networks. In the empirical literature, it is relatively standard to define α_i not only as her own observable characteristics but also as the observable *average* characteristics of individual i's neighbors (*contextual effects*) (Manski, 1993). As a result, α_i can be written as:

$$\alpha_i = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i} \sum_{m=1}^{M} \sum_{j=1}^{n} g_{ij}\, x_j^m \gamma_m \tag{17}$$

where $g_i = \sum_{j=1}^{n} g_{ij}$ is the number of direct links (the degree) of individual i, x_i^m is a set of M variables accounting for observable characteristics in individual characteristics of individual i, and β_m, γ_m are parameters. Using (17), equation (2) can be written as:

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + \sum_{m=1}^{M} \beta_m x_{i,r}^m + \frac{1}{g_{i,r}} \sum_{m=1}^{M} \sum_{j=1}^{n_r} g_{ij,r}\, x_{j,r}^m \gamma_m + \eta_r + \varepsilon_{i,r}. \tag{18}$$

[18]See also Campbell (2013) who develops a model based on random graphs for understanding the ways in which social learning, via word-of-mouth between friends, affects demand and, hence, the optimal pricing and advertising strategies of a firm.

This corresponds to the best-reply function of our model (and we know that there exists a unique Nash equilibrium in efforts if $\phi_1 < 1/\mu_1(\mathbf{G})$). There are at least three different econometric problems that need to be addressed in order to have a credible estimation of ϕ_1. First, there is the *reflection problem* (Manski, 1993), which is due to the difficulty in separating the *endogenous peer effect* (captured by ϕ_1) from the *contextual effect* (captured by the γ_m). Second, there is the *common-shock* problem, which is due to the fact that all individuals belonging to the same network r are affected by a common shock (for example, the teacher quality of a class) that affects their outcomes $y_{i,r}$. These two problems are usually solved by using the structure of the network (for example, one can use friends of friends' characteristics as an instrument for friends' actions) and network fixed effects η_r (Bramoullé, Djebbari and Fortin, 2009). Finally, there is the *correlated effect* so that individuals in the same network may behave similarly because they have similar unobserved individual characteristics. This issue is more difficult to address and researchers have either been explicitly modeling the network formation process (see, e.g. Mele, 2013; Goldsmith-Pinkham and Imbens, 2013; Chandrasekhar and Jackson, 2014; Badev, 2014), or using instrumental variables (Bifulco, Fletcher and Ross, 2011; Patacchini and Zenou, 2014) or natural experiments where agents are randomly allocated to the network (Carrell, Sacerdote and West, 2013; Algan et al., 2015; Hahn et al., 2015; Lindquist, Sauermann and Zenou, 2015). We refer to the literature surveys provided by Blume et al. (2011), Advani and Malde (2014), Jackson (2014), Jackson, Rogers and Zenou (2015), Graham (2015), Topa and Zenou (2015) and the chapter in this *Handbook* by Fortin and Boucher (2015) for a detailed treatment of these econometric issues.

3.2 Key players in criminal networks

"Key players" type of policies have been implemented in different countries. It is indeed estimated that only a few offenders are responsible for a very large proportion of all crimes committed. An example of such a policy is the *Operation Ceasefire* in the United States. In this type of operation, police beefed up patrols in the area, attempting to locate gang members who had *outstanding arrest warrants* or *had violated probation or parole regulations.* Gang members who had *violated public housing rules*, *failed to pay child support*, or were similarly vulnerable were also subjected to stringent enforcement (Tita et al., 2003). This latter policy combines a strong law enforcement response with a "pulling levers" deterrence effort aimed at chronic gang offenders. The key to the success is to use a "lever pulling" approach, which is a crime deterrence strategy that attempts to prevent violent behavior by using a *targeted individual or a group's vulnerability* to law enforcement as a means of gaining their compliance. Operation Ceasefire was first launched in Boston and youth homicide fell by two-thirds after the Ceasefire strategy was put in place in 1996 (Kennedy, 1998). It was then implemented in Los Angeles in 2000 and the results were very good. There was also a considerable decrease in crime.

There have been similar policies in the UK with a large scale policy intervention—the Street Crime Initiative (SCI)—that was introduced in England and Wales in 2002. This policy allocated additional resources to some police force areas to *specifically target street crime*, whereas other forces did not receive any additional funding. Machin and Marie (2011) show that robberies fell significantly in SCI police forces relative to non-SCI forces after the initiative had been introduced. Moreover, the policy seems to have been a cost effective one, even after extensively testing for possible displacement or diffusion effects on other crimes and into adjacent areas. Overall, they reach the conclusion that increased police resources can be used to generate falls in crime, at least in the context of the SCI program they study.

Machin, Marie and Priks (2014) investigate the impact of the targeting of the *most prolific offenders* on area crime rates. They empirically study reforms in England and Wales (between 2000 and 2004) and Sweden (introduced in 2012) where the police has shifted the focus from general policing towards prolific offenders. For this purpose, the police has constructed lists with names of prolific offenders and has allocated resources to closely monitoring these individuals (by, e.g., visiting offenders' homes). The reforms were cost-neutral, i.e., the government provided no extra funding. Using a difference-in-difference approach, they show that the policy was successful by reducing burglaries by approximately 5-10 percent in England, Wales and Sweden.

However, these policies are implemented based on intuitive judgements on whom to target rather than using a network approach. In particular, most policies have been targeting the most active or the most prolific criminals. We would now like to expose the empirical test of the key-player policy exposed in Section 2.2 and show how it outperforms more "intuitive" policies such as targeting the most active criminals.

Liu et al. (2012) were the first to test the key-player policy using the AddHealth data[19] for the *local-aggregate* model where the utility is given by (1). First, they estimate equation (18) and obtain an estimated value of ϕ_1 equal to 0.0457. Then, using this estimated value, they can calculate the key player for each network using the intercentrality measure (8). They find that the key player is *not* necessarily the most active criminal in the network. They

[19] The National Longitudinal Survey of Adolescent Health (Add Health) has been designed to study the impact of the social environment on adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in the years 1994-95. The most interesting aspect of the Add Health data is the information on friendships that is based on actual friend nominations and helps us create the friendship network.

also find that it is *not* straightforward to determine which delinquent should be removed from a network by only observing his or her criminal activities or position in the network. Compared to other criminals, the key players are less likely to be a female, are less religious, belong to families whose parents are less educated and have the perception of being more socially excluded. They also feel that their parents care less about them, are more likely to come from single-parent families and have more troubles getting along with their teachers. Finally, Liu et al. (2012) show that the key player policy outperforms other reasonable policies like targeting the most active criminals in the network. They calculate the average crime reduction for all 145 networks when a key-player policy, a random-target policy and a policy that removes the most active criminal are implemented. The authors have put together networks of the same size and calculated the average crime reduction for this network size under the three policies. For example, for all networks of size 4, the average crime reduction is 29.94 percent on average when the key-player policy is implemented, it is 23.86 percent for the random-target policy and around 25 percent when the most-active criminal policy is implemented.

Liu et al. (2014) test the key-player policy also with AddHealth data but for the general model with utility (14). We may first ask which model that best matches the data at hand: the local-aggregate model (utility given by (1)) or the local-average model (utility given by (12))? Liu, Patacchini and Zenou (2014) provide a test (the J test) that determines which model is more adequate for the data at hand. Liu et al. (2014) perform such a test and show that both models match the delinquency data of the AddHealth data well. As a result, they use the general model where the utility of each agent is given by (14). They also find that the key player is not necessarily the most *active* student in delinquent activities. Only in 19 networks out of the 103 networks in the sample is the key player the most active student in delinquent activities. They also look at crime reduction. They find that, for a network with 4 students (which corresponds to the median network in their sample), the percentage reductions in aggregate delinquency by removing the key player, the most active delinquent, and a random delinquent are 70.19%, 53.13%, and 50.71%, respectively. Hence, targeting key players is more effective than targeting the most active delinquent in reducing total delinquency.

Lindquist and Zenou (2014) also test the key-player policy but with a different dataset. They look at individuals in Sweden who are aged above 16 and who have been suspected of at least one crime. For this purpose, they have access to the official police register of all individuals who are suspected of committing a crime in Sweden. In this register, the police keeps records of who is suspected of committing a crime with whom. In this context, a (criminal) link exists between two individuals if they are suspected of committing a crime together (and are then convicted). Both the convictions data and the suspects data include crime type, crime date, and the sanction received. One advantage of this dataset over the AddHealth one is that links are not self-reported and are thus less subject to measurement errors. Another advantage is that information on links is available at each moment in time over a period of 20 years. As a result, they can add individual lagged crime as one of the individual level control variables.

They find an estimate of ϕ_1 of 0.167. This means that having only one friend increases own crime by 20 percent. If we consider the case of four individuals (their smallest network), then individual crime will increase by 100 percent compared to the case when the individual is committing a crime by herself. Lindquist and Zenou (2014) then consider two periods of three years each (2000 to 2002 and 2003 to 2005). The Period 1 data set includes 15,230 co-offenders who are suspected of committing (on average) 5.91 crimes each and who are distributed over 1,192 separate networks. The Period 2 data set includes 15,143 co-offenders who are suspected of committing (on average) 5.92 crimes each and who are distributed over 1,185 networks. Their data also include 3,881 individuals who are members of a network with four or more individuals in *both* periods. They show that 23% of all key players are not the most active criminal in their own networks; 23% do not have the highest eigenvector centrality; and 20% do not have the highest betweenness centrality.

Because they have two periods of time, Lindquist and Zenou (2014) can test the prediction of crime reduction following the key-player policy against the true outcome observed in Period 2 data. They thus look at the relative effect of removing the key player in those cases in which the key player is no longer part of the active network. For this purpose, they create an indicator variable for each person indicating whether or not they have died during the relevant time period and whether they have been placed in prison. Their results indicate that, in the real-world, the key player policy outperforms the random player policy by 9.58%. The key player policy also outperforms the policy of removing the most active player by 3.16% and the policy of removing the player with the highest eigenvector and betweenness centrality by 8.12% and 2.09%, respectively.

3.3 Key players in R&D networks

R&D partnerships have become a widespread phenomenon characterizing technological dynamics, especially in industries with a rapid technological development such as, for instance, the pharmaceutical, chemical and computer industries (see, e.g. Hagedoorn, 2002; Powell et al., 2005). In those industries, firms have become more specialized in specific domains of a technology and they tend to combine their knowledge with the knowledge of other firms that are specialized in different technological domains (Powell et al., 1996; Weitzman, 1998). The increasing importance of R&D collaborations has spurred research for theoretical models studying these relationships, and for empirical tests of these models.

We would like to determine the key players or, more exactly, the key firms in R&D networks. In other words, we would like to determine which firms that are crucial for an industry in the sense that, if they exit the market (i.e. go bankrupt), the cost in terms of total activity or welfare for the remaining firms and consumers will be the highest possible. Following König, Liu and Zenou (2014), consider the above model where the utility (or profit) function of each firm i is similar to (9). They estimate equation (11) and obtain values for both ϕ_1 and ρ with the predicted signs so that there are positive spillover effects of R&D collaborations, captured by $\phi_1 > 0$ and negative competition effects, captured by $-\rho < 0$. Using the MERIT-CATI dataset, they then determined the key firms over a period of more than 40 years . König, Liu and Zenou (2014) show that the key firms are usually not those with the largest number of R&D collaborations (degree), the largest number of patents, nor the highest eigenvector, betweenness or closeness centrality and, more importantly, not the firm with the highest market share in its sector. Interestingly, General Motors, which was bailed out in 2008 by President-elected Obama, was among the key firms. They show that, if General Motors had been removed from the market in 1990, then total welfare would have been reduced by 8.37 % while total output would have been decreased by 2.14 %.

3.4 Key players in education

The influence of peers on education outcomes has been widely studied both in economics and sociology (Sacerdote, 2011). There are, however, fewer papers studying the network effects on education (see, in particular, Calvó-Armengol, Patacchini and Zenou, 2009; Bifulco, Fletcher and Ross, 2011; Patacchini, Rainone and Zenou, 2014). Using a model similar to that of Section 2.2.1 and the AddHealth data, Calvó-Armengol, Patacchini and Zenou (2009) were able to show that the Katz-Bonacich centrality of a student is a key determinant of his/her grade. Here, we would like to go further by examining the importance of different centralities, including the Katz-Bonacich and the key-player centrality defined in (8), on educational outcomes.

In Section 3.1, we have seen that one of the most difficult empirical challenges in the estimation of network effects was the endogeneity of the adjacency matrix G or, equivalently, the fact that network formation is endogenous and may be due to correlation in unobservables between the individuals involved in the link formation. One way of addressing this issue is to consider a random experiment that exogenously assigns links to individuals. This is what is performed by Hahn et al. (2015), who propose to test the importance of centrality measures in educational achievement. They implement a field experiment in Bangladesh, whose design involved randomly grouping students within-classroom among *grade-four students* in rural primary schools. These experiments were conducted in 80 schools in two districts (Khulna and Satkhira) in Bangladesh.

Let us now describe the experiment. In June 2013, the authors conducted a survey of all students in the 80 schools by asking them to name up to ten closest friends (who defined the network, as in the AddHealth data), starting from the most to the least close friends. They also conducted a separate household survey, which contained questions on parent education, parent age, parent occupation, and other household characteristics and asked each student to perform a *baseline math test* to measure her ability, which will help the authors balance groups by average ability. In July 2013, one month later, groups were randomly allocated. Each group contains four students and any student has an equal chance of being in one group or the other since groups are generated by a random-number generator. To implement random grouping that has a relatively similar mean across groups, they rank students by their baseline test score. Hahn et al. (2015) then randomly select a student from each quartile to form a group of size four. Then, newly (randomly) formed groups were asked to solve a *general knowledge test*, which took place in July 2013 and which is performed *collectively* by each group of four. Each group was given a *math assignment* to be completed *collectively* by the end of the week. Finally, after each group had handed over its math assignment, each individual had to take an *individual math test*. Prizes were given to the most successful students.

Hahn et al. (2015) want to investigate the following question: *If, by chance, a student ends up in a group with a high average centrality, does this positively impact on his/her test score as compared to someone who finds him/herself in a group with a lower average centrality?* They consider six measures of centrality: the degree, closeness, betweenness, eigenvector, Katz-Bonacich and key-player (or intercentrality) centrality. The authors calculate the average centrality of each group of four (not including i's centrality) and compare them across the classroom (there are usually 40 students per classroom and thus 10 groups of 4) and across the 80 schools where groups were randomized.

They find that the higher is the average centrality of a group to which a student i is allocated, the higher is his/her grade. This is true for both the *general knowledge test* (a homework performed by the group) and the *individual math test* (performed individually). This result is also true for *any* centrality measure with different magnitudes of the effects.

Since these centrality measures are correlated, the authors then test one measure against the other. Interestingly, they show that for both the general knowledge test and the individual math test, the (average) *Katz-Bonacic*h centrality as well as the key-player centrality have the most significant impact on educational outcomes. They also test the role of leadership in a group on own grade outcomes by looking at the impact of the individual with the highest centrality in the group on own outcome. They find that, for both tests, it is the Katz-Bonacich centrality that performs best. These results indicate that the composition of a group in terms of centrality is of importance for both individual and group outcomes and that Katz-Bonacich and key–player centralities are key for these outcomes.

3.5 Key players in financial networks

Since the onset of the financial crisis in 2007, the discourse regarding bank safety has shifted strongly from the riskiness of financial institutions as individual firms to concerns about systemic risk. As the crisis evolved, so did the public debate, with concerns about *systemic risk* evolving from too-big-to-fail (TBTF) considerations to too-interconnected-to-fail (TITF) ones. As a result, it seems quite natural to consider the issue of key players or key banks in financial networks. Systemic risk is usually defined as the risk of default of a large portion of the financial system, which depends on the network of financial exposures among institutions.

There are several recent papers that model financial networks using network theory (see Acemoglu, Ozdaglar and Tahbaz-Salehi, 2015a; Elliott, Golub and Jackson, 2014; Cohen-Cole, Patacchini and Zenou, 2015; just to mention a few) but do not analyze the targeting of banks in order to reduce the risk of contagion.[20] Demange (2014) has proposed a *threat index*, which measures the decrease in payment within the banking system following a reduction in net worth at one institution. The latter is related to the intercentrality measure of the key player. Battiston et al. (2012) propose another index called the *DebtRank*, which is a novel measure of systemic impact inspired by feedback-centrality that takes recursively into account the impact of the distress of an initial node across the whole network. Battiston et al. (2012) apply their methodology to analyze a dataset on the USD 1.2 trillion FED emergency loans program to global financial institutions during 2008-2010. They find that a group of 22 institutions, which received most of the funds, forms a strongly connected graph where each of the nodes becomes systemically important at the peak of the crisis.

[20]See the chapter in this Handbook by Acemoglu, Ozdaglar and Tahbaz-Salehi (2015b) for an overview of the literature on systemic risk and networks and the one by Cabrales, Gale and Gottardi (2015) for an overview of the literature on financial contagion.

Finally, Denbee et al. (2014) slightly modify the benchmark model of Section 2.2 to model banks' liquidity holding decision as a simultaneous game on an interbank borrowing network. They consider a network between banks which is the sterling unsecured overnight interbank money market. This is where banks lend central bank reserves to each other, unsecured, for repayment the following day. The strength of the link between any two banks in their network is measured using the fraction of borrowing by one bank from the other. Hence, their network is weighted and directional. As well as relying on their own liquidity buffers, banks can also rely on their borrowing relationship within the network to meet unexpected liquidity shocks. Using daily data from January 2006 to September 2010, the authors estimate model (18) and find evidence for a substantial, and time varying, network risk. In the period before the Lehman crisis, the network is cohesive and liquidity holding decisions are complementary and there is a large network liquidity multiplier. During the 2007-08 crisis, the network becomes less clustered and liquidity holding less dependent on the network. After the crisis, during Quantitative Easing, the network liquidity multiplier becomes negative, implying a lower network potential for generating liquidity.

They also identify *risk key players*, that is, the banks that contribute the most to the aggregate liquidity risk through these three periods. They show that the risk that key players took during these periods varies a great deal. They also find that the key players in the network are not necessarily the largest borrowers. In fact, during the credit boom, large lenders and borrowers are equally likely to be key players. This set of findings is of policy relevance, and gives guidance on how to effectively inject liquidity, to reduce the network risk, if the government decides to intervene.

3.6 Key players and diffusion in networks

Banerjee et al. (2013) structurally estimate the model developed in Section 2.2.3 to the diffusion of microfinance loans in India, in a setting where the set of potentially first-informed individuals is known. They show that the *communication centrality* of the injection points is a strong predictor of eventual participation in microfinance and should therefore provide guidance to anyone trying to spread the news about microfinance in similar villages. They define communication centrality as follows. For each leader (injection points), they compute a score. This score is the fraction of households that would eventually participate if this household were the only one initially informed. To compute this fraction, they simulate the model with information passing and participation decisions being governed by the estimated values of ϕ_1^N, ϕ_1^P, and β. Banerjee et al. (2013) call this score the communication centrality of a node. However, the communication centrality cannot be computed without the estimations of ϕ_1^N, ϕ_1^P, and β, which could be very different if we were interested in the diffusion of other products or even microfinance in a very different context. As a result, the authors propose an approximation of communication centrality, called *diffusion centrality*, which is highly correlated with communication centrality, but requires considerably less data. In particular, it does not rely on estimating the diffusion model. The diffusion centrality of a node i in a network with an adjacency matrix $\mathbf{G}$, passing probability $\phi_1^N = \phi_1^P = \phi_1$, and iterations T, as the ith entry of the following vector:

$$\boldsymbol{\delta}\left(\mathbf{g}, \phi_1, T\right) = \left[\sum_{t=1}^{T} \left(\phi_1 \, \mathbf{G}\right)^t\right] \mathbf{1} \tag{19}$$

Interestingly, the diffusion centrality of a node i becomes proportional to either Katz-Bonacich centrality or eigenvector centrality when $T \to +\infty$, depending on whether ϕ_1 is smaller than the inverse of the largest eigenvalue of the adjacency matrix or exceeds it (see Proposition 1). In the intermediate region of T, the measure differs from existing measures. As for the Katz-Bonacich centrality defined in (4) and the intercentrality or key-player centrality defined in (7), any method for computing a measure of diffusion centrality relies on the estimation of ϕ_1 or the choice of an appropriate value for ϕ_1. Extreme values of ϕ_1 either lead to no diffusion or to complete diffusion and thus, do not distinguish nodes. Barnerjee et al. (2013) choose a prominent intermediate value of ϕ_1: the inverse of the largest eigenvalue of the adjacency matrix, $\mu_1(\mathbf{G})$. This is the critical value of ϕ_1 for which the entries of $\left(\phi_1 \, \mathbf{G}\right)^T$ tend to 0 as T grows if $\phi_1 < 1/\mu_1(\mathbf{G})$ and some entries diverge if $\phi_1 > \mu_1(\mathbf{G})$.

Another interesting related application of the key-player policy (i.e. targeting injection points) is by Banerjee et al. (2014) where they try to identify the key player (eigenvector centrality, diffusion centrality) without knowing the whole network but by asking people about the central players. Indeed, in many instances, learning who is central in a social network has the potential of being difficult and costly. Even for members of the community, knowing the structure of the network beyond their immediate friends is far from straightforward. As a result, the authors would like to answer the following question: Can we identify the members of a community who are best-placed to diffuse information simply by asking a random sample of individuals? Using the same data on 35 Indian villages as in Banerjee et al. (2013), they show that, by tracking sources of gossip, one can identify those who are most central in a network according to the "diffusion centrality," defined in (19). In particular, the authors find that respondents *accurately* nominate those who are diffusion central (not just those with many friends). In other words, they find that individuals in a network are able to identify central individuals within their community without knowing anything about the structure of the network.

4 Concluding remarks

We provide an overview of the literature on key players in networks. There are different measures in the literature, which fundamentally capture different aspects of centrality and are therefore related to different behaviors. We have seen that both eigenvector and Katz-Bonacich centralities are crucial in explaining educational outcomes and diffusion processes. This is also true in other contexts. For example, Acemoglu et al. (2012) focus on production networks and show that the sector with the highest *influence vectors*, a measure closely related to the Katz-Bonacich centrality, is the one with the highest share in total output and is more important to cause aggregate fluctuations. In other words, sectors that take more "central" positions in the network representation of the economy play a more important role in determining aggregate output. Studying the direct and spillover effects of local state capacity using the network of Colombian municipalities, Acemoglu, Garcia-Jimeno and Robinson (2015) find that Katz-Bonacich centrality, betweenness centrality and local clustering are strong predictors of where more state capacity should be directed.

In this chapter, we introduce a new centrality measure, intercentrality, which determines the key player, who is the agent targeted by the planner so that, once removed, she generates the highest level of reduction in total activity. This centrality measure is different to those proposed in the literature since it has a normative aspect. We believe that this key-player policy based on this new measure makes sense when there are strategic complementarities in actions between the different agents in a network such as, for example, in crime, education, R&D collaborations, interbank loans, etc. We also consider another notion of key player based on targeting agents that are optimally positioned to quickly diffuse information, attitudes, behaviors or goods. Then, we examine the empirical tests of the key-player policies for criminal networks, education, R&D networks, financial networks and diffusion of microfinance. We show that implementing such a policy outperforms other standard policies such as targeting the most active agents in a network.

We believe that key-player policies are crucial when resources are limited and multiplier effects are important. We have seen how these policies can be applied to different aspects of the economy, such as crime or education, but we believe that they have broader implications. For example, König et al. (2014) study the key-player policies for wars. They look at the recent civil war in the Democratic Republic of Congo (DRC), which involves many groups and a complex network of alliances and rivalries. They examine how the removal of each group involved in the conflict would reduce the conflict intensity. They show that, while large groups are, on average, more crucial than small ones, the relationship is not one-to-one. The hypothetical removal of some relatively small players turns out to have large effects on the containment of the DRC conflict. We hope that more empirical studies will be implemented in the future showing the relevance of key players in many other aspects of the economy.

References

[1] Acemoglu, D., Carvalho, V.M., Ozdaglar, A. and A. Tahbaz-Salehi (2012), "The network origins of aggregate fluctuations," *Econometrica* 80, 1977-2016.

[2] Acemoglu, D., Ozdaglar, A. and A. Tahbaz-Salehi (2015a), "Systemic risk and stability in financial networks," *American Economic Review* 105, 564-608.

[3] Acemoglu, D., Ozdaglar, A. and A. Tahbaz-Salehi (2015b), "Systemic risk and networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[4] Acemoglu, D., Garcia-Jimeno, C. and J.A. Robinson (2015), "State capacity and economic development: A network approach," *American Economic Review*, forthcoming.

[5] Advani, A. and B. Malde (2014), "Empirical methods for networks data: Social effects, network formation and measurement error," Unpublished manuscript, University College London.

[6] Algan, Y., Do, Q.-A., Le Chapelain, A. and Y. Zenou (2015), "How do social networks shape our values? A natural experiment among future French politicians„" Unpublished manuscript, Sciences Po, Paris.

[7] Badev, A. (2014), "Discrete Games in endogenous networks: Theory and policy," Unpublished manuscript, Federal Reserve Board, Washington D.C.

[8] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: the key player," *Econometrica* 74, 1403-1417.

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

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

[11] Banerjee, A., Chandrasekhar, A.G., Duflo, E. and M.O. Jackson (2013), "The diffusion of microfinance," *Science* 341, 6144.

[12] Banerjee, A., Chandrasekhar, A.G., Duflo, E. and M.O. Jackson (2014), "Gossip: Identifying central people in social networks," NBER Working Paper No. 20422.

[13] Battiston, S., Puliga, M., Kaushik, R., Tasca, P. and G. Caldarelli (2012), "DebtRank: Too central to fail? Financial networks, the FED and systemic risk," *Scientific Reports* 2, 541.

[14] Belhaj, M., Bervoets, S. and F. Deroïan (2014), "Using network design as a public policy tool," Unpublished manuscript, Aix-Marseille School of Economics.

[15] Bifulco, R., Fletcher, J.M. and S.L. Ross (2011), "The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health," *American Economic Journal: Economic Policy* 3, 25-53.

[16] Bloch, F. (2015), "Pricing and advertising in networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[17] Blume, L.E., Brock, W.A., Durlauf, S.N. and Y.M. Ioannides (2011), "Identification of social interactions," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science.

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

[19] Borgatti, S.P. (2003), "The key player problem," In: R. Breiger, K. Carley and P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, New York: National Academy of Sciences Press, pp. 241-252.

[20] Borgatti, S.P. (2006), "Identifying sets of key players in a network," *Computational, Mathematical and Organizational Theory* 12, 21-34.

[21] Bramoullé, Y., Djebbari, H. and B. Fortin (2009), "Identification of peer effects through social networks," *Journal of Econometrics* 150, 41-55.

[22] Bramoullé, Y. and R. Kranton (2007), "Public goods in networks," *Journal of Economic Theory* 135, 478-494.

[23] Bramoullé, Y. and R. Kranton (2015), "Network games," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[24] Bramoullé, Y., Kranton, R. and M. D'Amours (2014), "Strategic interaction and networks," *American Economic Review* 104, 898-930.

[25] Burt, R.S. (1992), *Structural Holes: The Social Structure of Competition*. Harvard: Harvard University Press.

[26] Cabrales, A., Gale, D. and P. Gottardi (2015), "Financial contagion," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[27] Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009), "Peer effects and social networks in education," *Review of Economic Studies* 76, 1239-1267.

[28] Calvó-Armengol, A. and Y. Zenou (2004), "Social networks and crime decisions. The role of social structure in facilitating delinquent behavior," *International Economic Review* 45, 939-958.

[29] Campbell, A. (2013), "Word-of-mouth communication and percolation in social networks," *American Economic Review* 103, 2466-2498.

[30] Carrell, S.E., Sacerdote, B.I. and J.E. West (2013), "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica* 81, 855-882.

[31] Chandrasekhar, A. and M.O. Jackson (2014), "Tractable and consistent random graph models," NBER Working Paper No. 20276.

[32] Chen, Y.-J., Zenou, Y. and J. Zhou (2015), "Multiple activities for socially connected criminals," Unpublished university, Stockholm University.

[33] Cohen-Cole, E., Patacchini, E. and Y. Zenou (2015), "Static and dynamic networks in interbank markets," *Network Science*, forthcoming.

[34] Dequiedt, V. and Y. Zenou (2014), "Local and consistent centrality measures in networks", CEPR Discussion Paper No. 10031.

[35] Demange, G. (2014), "Contagion in financial networks: A threat index," Unpublished manuscript, Paris School of Economics.

[36] Denbee, E., Julliard, C., Li, Y. and K. Yuan (2014), "Network risk and key players: A structural analysis of interbank liquidity," Unpublished manuscript, London School of Economics and Political Science.

[37] de Marti, J. and Y. Zenou (2014), "Networks games under incomplete information," CEPR Discussion Paper No. 10290.

[38] Domingos, P. and M. Richardson (2001), "Mining the network value of customers," Proceedings of the 7th International Conference on Knowledge Discovery and Data Mining, pp. 57-66.

[39] Elliott, M.L., Golub, B. and M.O. Jackson (2014), "Financial networks and contagion," *American Economic Review* 104, 3115-3153.

[40] Fortin, B. and V. Boucher (2015), "Some challenges in the empirics of the effects of networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[41] Galeotti, A. and S. Goyal (2009), "Influencing the influencers: A theory of strategic diffusion," *Rand Journal of Economics* 5, 1-32.

[42] Galeotti, A. and B. Rogers (2013), "Strategic immunization and group structure," *American Economic Journal: Microeconomics* 40, 509-532.

[43] Glaeser, E.L. Sacerdote, B. and J. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111, 508-548.

[44] Goldsmith-Pinkham, P. and G.W. Imbens (2013), "Social networks and the identification of peer effects," *Journal of Business and Economic Statistics* 31, 253-264.

[45] Goyal, S. (2007), *Connections: An Introduction to the Economics of Networks*, Princeton: Princeton University Press.

[46] Graham, B.S. (2015), "Methods of identification in social networks," *Annual Review of Economics* 7, forthcoming.

[47] Hagedoorn, J. (2002), "Inter-firm R&D partnerships: An overview of major trends and patterns since 1960," *Research Policy* 31, 477-492.

[48] Hahn, Y., Islam, A., Patacchini, E. and Y. Zenou (2015), "Network and peer effects in education: Evidence from a field experiment in Bangladesh," Unpublished manuscript, Stockholm University.

[49] Hartline, J., Mirrokni, V.S. and M. Sundararajan (2008), "Optimal marketing strategies over social networks," International World Wide Web Conference Committee (IW3C2), Beijing, China.

[50] Hiller, T. (2014), "Peer effects in endogenous networks," Unpublished manuscript, University of Bristol.

[51] Husslage, B., Borm, P., Burg, T., Hamers, H. and R. Lindelauf (2014), "Ranking terrorists in networks: A sensitivity analysis of Al Qaeda's 9/11 attack," CentER Discussion Paper No. 2014-028, Tilburg University.

[52] Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

[53] Jackson M.O. (2008), *Social and Economic Networks*, Princeton, NJ: Princeton University Press.

[54] Jackson, M.O. (2011), "An overview of social networks and economic applications," In: J. Benhabib, A. Bisin and M.O. Jackson (Eds.), *Handbook of Social Economics Volume 1A*, Amsterdam: Elsevier Science, pp. 511-579.

[55] Jackson, M.O. (2014), "Networks in the understanding of economic behaviors," *Journal of Economic Perspectives* 28, 3-22.

[56] Jackson, M.O., Rogers, B.W. and Y. Zenou (2015), "The impact of social networks on economic behavior," Unpublished manuscript, Stanford University.

[57] Jackson, M.O. and L. Yariv (2011), "Diffusion, strategic interaction, and social structure," In: J. Benhabib, A. Bisin and M.O. Jackson (Eds.), *Handbook of Social Economics Volume 1A*, Amsterdam: Elsevier Science, 645-678.

[58] Jackson, M.O. and Y. Zenou (2014), "Games on networks", In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier Publisher, pp. 91-157.

[59] Kahan, D.M. (1997), "Social influence, social meaning, and deterrence," *Virginia Law Review* 83, 349-395

[60] Katz, L. (1953), "A new status index derived from sociometric analysis," *Psychometrika* 18, 39-43.

[61] Kempe, D., Kleinberg, J. and É. Tardos (2003), "Maximizing the spread of influence in a social network," Proceedings of the 9th International Conference on Knowledge Discovery and Data Mining, pp. 137-146.

[62] Kempe, D., Kleinberg, J. and É. Tardos (2005), "Influential nodes in a diffusion model for social networks," Proceedings of the 32nd International Colloque on Automata, Languages and Programming, pp. 1127-1138.

[63] Kennedy, D.M. (1998), "Pulling levers: Getting deterrence right," *National Institute of Justice Journal* 236, 2-8.

[64] Kleiman, M.A. (2009), *When Brute Force Fails. How to Have Less Crime and Less Punishment*, Princeton: Princeton University Press.

[65] Kleinberg, J. (2007), "Cascading behavior in networks: Algorithmic and economic issues," In: N. Nisan, T. Roughgarden, É. Tardos and V. Vazirani (Eds.), *Algorithmic Game Theory*, Cambridge: Cambridge University Press, pp. 613-632.

[66] König, M.D., Tessone, C. and Y. Zenou (2014), "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics* 9, 695-752.

[67] König, M.D., Liu, X. and Y. Zenou (2014), "R&D networks: Theory, empirics and policy implications," CEPR Discussion Paper No. 9872.

[68] König, M.D., Rohmer, D., Thoenig, M. and F. Zilibotti (2015), "Networks in conflict: Theory and evidence from the great war of Africa," Unpublished manuscript, University of Zurich.

[69] Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D. and and O. Zlotowski (2005), "Centrality indices," In: Brandes, U. and T. Erlebach, T (Eds.), *Network Analysis: Methodological Foundations*, Lecture Notes in Computer Science No.. 3418, New York: Springer-Verlag, pp. 16-61.

[70] Lagerås, A. and D. Seim (2014), "Strategic complementarities, network games and endogenous network formation," Unpublished manuscript, Stockholm University.

[71] Lamberson, P.J. (2015), "Diffusion in networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[72] Lindelauf, R., Hamers, H. and B. Husslage (2013), "Cooperative game theoretic centrality analysis of terrorist networks: The cases of Jemaah Islamiyah and Al Qaeda," *European Journal of Operational Research* 229, 230-238.

[73] Lindquist, M.J., Sauermann, J. and Y. Zenou (2015), "Network effects on worker productivity and on-the-job-training in the Netherlands," Unpublished manuscript, Stockholm University.

[74] Lindquist, M.J. and Y. Zenou (2014), "Key players in co-offending networks," CEPR Discussion Paper No. 9889.

[75] Liu, X., Patacchini, E. and Y. Zenou (2014), "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization* 103, 39-59.

[76] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper No. 8772.

[77] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2014), "Who is the key player? A network analysis of juvenile delinquency," Unpublished manuscript, Stockholm University.

[78] Mahadev, N.V.R. and U.N. Peled (1995), *Threshold Graphs and Related Topics*, Amsterdam: North Holland.

[79] Machin, S. and O. Marie (2011), "Crime and police resources: The street crime initiative," *Journal of the European Economic Association* 9, 678-701.

[80] Machin, S., Marie, O. and M. Priks (2014), "Targeting prolific offenders to reduce crime: Theory and evidence from two European experiments," Unpublished manuscript, Stockholm University.

[81] Manski, C.F. (1993), "Identification of endogenous effects: The reflection problem," *Review of Economic Studies* 60, 531-542.

[82] Mayzlin, D. (2015), "Marketing and networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, forthcoming.

[83] Mele, A. (2013), "A structural model of segregation in social networks," Unpublished manuscript, Johns Hopkins University, Carey Business School.

[84] Ortiz-Arroyo, D. (2010), "Discovering sets of key players in social networks," In: A. Abraham, A.-E., Hassanien and V. Snášel (Eds.), *Computational Social Network Analysis*, London: Springer Verlag, pp. 27-47.

[85] Patacchini, E., Rainone, E. and Y. Zenou (2014), "Heterogeneous peer effects in education," CEPR Discussion Paper No. 9804

[86] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economics, and Organization* 28, 1-31.

[87] Patacchini, E. and Y. Zenou (2014), "Social networks and parental behavior in the intergenerational transmission of religion", Unpublished manuscript, Stockholm University.

[88] Powell, W.W., Koput, K.W. and L. Smith-Doerr (1996), "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology," *Administrative Science Quarterly* 41, 116-145.

[89] Powell, W.W., White, D.R., Koput, K.W. and J. Owen-Smith (2005), "Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences,"*American Journal of Sociology* 110, 1132-1205.

[90] Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of Economics of Education, Vol. 3*, Amsterdam: Elevier Science, pp. 249-277.

[91] Sathik, M. and A. Rasheed (2009), "A centrality approach to identify sets of key players in an online weblog," *International Journal of Recent Trends in Engineering* 2, 85-87.

[92] Silverman, D. (2004), "Street crime and street culture," *International Economic Review* 45, 761-786.

[93] Sommer, M. (2014), "Centrality with vertex idiosyncracy: Comparative statics and vertex-weighted key-player analysis," Unpublished manuscript, Stockholm University.

[94] Tita, G.K. Riley, J., Ridgeway, G., Grammich, C., Abrahamse, A. and P.W. Greenwood (2003), *Reducing Gun Violence: Results from an Intervention in East Los Angeles*, Santa Monica, CA: RAND Corporation.

[95] Topa, G. and Y. Zenou (2015), "Neighborhood and network effects", In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5*, Amsterdam: Elsevier Publisher, forthcoming.

[96] Verdier, T. and Y. Zenou (2004), "Racial beliefs, location and the causes of crime," *International Economic Review* 45, 731-760.

[97] Wilson, J.Q. and R.J. Herrnstein (1985), *Crime and Human Nature*, New York: Simon and Schuster.

[98] Wasserman, S. and K. Faust (1994), *Social Network Analysis. Methods and Applications*, Cambridge: Cambridge University Press.

[99] Weitzman, M.L. (1998), "Recombinant growth," *Quarterly Journal of Economics* 113, 331-360.

[100] Zenou, Y. (2015), "Networks in economics," In: J.D. Wright (Ed.), *International Encyclopedia of Social and Behavioral Sciences, 2nd Edition*, Amsterdam: Elsevier Publisher, forthcoming.