Juvenile Delinquency and Conformism*

Eleonora Patacchini† Yves Zenou‡

May 13, 2008

Abstract

This paper studies whether conformism behavior affects individual outcomes in crime. We present a social network model of peer effects with ex-ante heterogeneous agents and show how conformism and deterrence affect criminal activities. We then bring the model to the data by using a very detailed dataset of adolescent friendship networks. A novel social network-based empirical strategy allows us to identify peer effects for different types of crimes. We find that conformity plays an important role for all crimes, especially for petty crimes. This suggests that, for juvenile crime, an effective policy should not only be measured by the possible crime reduction it implies but also by the group interactions it engenders.

Key words: social networks, linear-in-means model, spatial autoregressive model, social norms.

JEL Classification: A14, C21, D85, K42, Z13.

*This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu).

†Università di Roma “La Sapienza”, Italy. E-mail: eleonora.patacchini@uniroma1.it.

‡Corresponding author. Stockholm University and Research Institute of Industrial Economics (IFN), Sweden. E-mail: yves.zenou@ne.su.se. Also affiliated with GAINS and CEPR.

1 Introduction

A large literature has developed on the general causes of, and the impact of public policy on, crime. Yet, no consensus has emerged on quite basic issues, such as, for example, the effects of incentives, both positive and negative, on crime.

Juvenile crime is an important aspect of this debate. According to the U.S. Department of Justice, juveniles were involved in 16 percent of all violent arrests and 32 percent of all property crime arrests in 1999. In addition, more than 100,000 juveniles are held in residential placement on any given day in the United States. However, despite these figures there are still many unanswered questions about juvenile crime. Some have shown that deterrence has a negative impact on juvenile crime (Levitt, 1998; Mocan and Rees, 2005;). It has also been shown that crime committed by younger people have higher degrees of social interactions (Glaeser et al., 1996; Jacob and Lefgren, 2003; Patacchini and Zenou, 2008).[1]

There is indeed a growing literature in economics suggesting that peer effects are very strong in criminal decisions. Case and Katz (1991), using data from the 1989 NBER survey of youths living in low-income Boston neighborhoods, find that the behaviors of neighborhood peers appear to substantially affect criminal activities of youth behaviors. They find that the direct effect of moving a youth with given family and personal characteristics to a neighborhood where 10 percent more of the youths are involved in crime than in his or her initial neighborhood is to raise the probability the youth will become involved in crime by 2.3 percent. Ludwig et al. (2001) and Kling et al. (2005) explore this last result by using data from the Moving to Opportunity (MTO) experiment that relocates families from high- to low-poverty neighborhoods. They find that this policy reduces juvenile arrests for violent offences by 30 to 50 percent of the arrest rate for control groups. This also suggests very strong social interactions in crime behaviors. Patacchini and Zenou (2008) test the role of weak ties in explaining criminal activities, revealing that weak ties have a statistically significant and positive effect on both the probability to commit crime and on its level. Finally, Bayer et al. (2007) consider the influence that juvenile offenders serving time in the same correctional facility have on each other's subsequent criminal behavior. They find strong evidence of peer effects in criminal activities since exposure to peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates with that crime.

The aim of the present paper is to analyze the role of conformism in juvenile crime

[1] In the crime literature, the positive correlation between self-reported delinquency and the number of delinquent friends reported by adolescents has proven to be among the strongest and one of the most consistently reported findings (see e.g. War, 1996, 2002; Matsueda and Anderson, 1998; Haynie, 2001).

using a network perspective. There are two important challenges in the empirics of social interactions: *(i)* the assessment of the *existence* of the endogenous effect of peers; *(ii)* the explanation of *how* peers influence each other, i.e. the mechanism generating such social interactions.[2]

We first present a social network model where individual utility depends on conformism. Conformism is the idea that the easiest and hence best life is attained by doing one's very best to blend in with one's surroundings, and to do nothing eccentric or out of the ordinary in any way. It may well be best expressed in the old saying, "When in Rome, do as the Romans do". To be more specific, using an explicit network analysis,[3] we develop a model where conformism[4] associated with deterrence are the key determinants of criminal activities. Our model is as follows. Each criminal belongs to a group of best friends and derives utility from exerting crime effort. We have a standard costs/benefits structure a la Becker with an added element, conformism. The new aspect of this model is that the social norm is endogenous and depends on the structure of the network. Indeed, direct friends define a social norm and depending of the location in the network, each individual has a different reference group. The utility function is such that each individual wants to minimize the social distance between his/her crime level and that of his/her reference group.

We derive the Nash equilibrium of this game and obtain that, when individuals are ex ante heterogenous (for example different race, sex, parents' education, etc.), they provide effort proportional to that of their reference group of best friends and that deterrence reduces crime. An interesting result is that, when individuals are ex ante identical, i.e. differ only by their location in the network, then, in equilibrium, all agents provide the same effort level. In other words, the Bonacich centrality index [5] is the same for all individuals in the network. This is a surprising result since Ballester et al. (2006), using a similar social network model

[2] See, in particular, the special issue on peer effects in the *Journal of Applied Econometrics* (Durlauf and Moffitt, 2003).

[3] There is a growing literature on networks in economics. See the recent literature survey by Jackson (2007).

[4] In economics, different aspects of conformism and social norms have been explored from a theoretical point of view. To name a few, (*i*) peer pressures and partnerships (Kandel and Lazear 1992) where peer pressure arises when individuals deviate from a well-established group norm, e.g., individuals are penalized for working less than the group norm, (*ii*) religion (Iannaccone 1992, Berman 2000) since praying is much more satisfying the more average participants there are, (*iii*) social status and social distance (Akerlof 1980, 1997, Bernheim 1994, Battu et al., 2007, among others) where deviations from the social norm (average action) imply a loss of reputation and status.

[5] To be more precise, the Bonacich centrality measure takes into account both direct and indirect friends of each individual but puts less weight to distant friends.

but without conformism, find that, when individuals are ex ante identical, each of them will provide a different effort level depending on his/her location in the network (as measured by his/her Bonacich index). Our result is due to the fact that the cost of deviating from the norm is sufficiently high so that individuals behave identically in equilibrium. However, when an additional heterogeneity is introduced (apart from the location of the network, individuals are heterogenous in their ability of committing crime, which is correlated with their idiosyncratic characteristics), individuals deviate from the social norm and behave partly according to their ability.

Even quite different, this theoretical model is along the lines of the growing literature on the social aspects of crime. In Sah (1991), the social setting affects the individual perception of the costs of crime, and is thus conducive to a higher or a lower sense of impunity. In Glaeser et al. (1996), criminal interconnections act as a social multiplier on aggregate crime. Calvó-Armengol and Zenou (2004) and Ballester et al. (2006) develop social network models of pure peer effects and no conformism.[6]

We then test our model using the U.S. National Longitudinal Survey of Adolescent Health (AddHealth), which contains unique detailed information on friendship relationships among delinquent teenagers. Empirical tests of models of social interactions are quite problematic because of well-known issues that render the identification and measurement of peer effects quite difficult: (i) reflection, which is a particular case of simultaneity (Manski, 1993) and (ii) endogeneity, which may arise for both peer self-selection and unobserved common (group) correlated effects.

In this paper, we exploit the architecture of social networks to overcome this set of problems and to achieve the identification of endogenous peer effects. More specifically, in social networks, each agent has a different peer-group, i.e. different friends with whom each teenager directly interacts. This feature of social networks guarantees the presence of excluded friends from the reference group (peer-group) of each agent, which are however included in the reference group of his/her best (direct) friends. Therefore, the exogenous characteristics of excluded friends are a natural set of instruments to overcome the reflection problem. This identification strategy is similar in spirit to the one used in the standard simultaneous equation model, where at least one exogenous variable needs to be excluded from each equation. In addition, because we observe individuals over networks, we can use a spec-

[6] Linking social interactions with crime has also been done in dynamic general equilibrium models (İmrohoroğlu et al., 2000, and Lochner 2004) and in search-theoretic frameworks (Burdett et al., 2003, 2004, and Huang et al., 2004). Other related contributions on the social aspects of crime include Silverman (2004), Verdier and Zenou (2004), Calvó-Armengol et al. (2007), Ferrer (2008).

ification of the empirical model with a network-specific component. By doing so, we are able to control for the presence of network-specific unobserved factors affecting both individual and peer behaviors. Such factors might be important omitted variables driving the sorting of agents into networks or effects arising from unobservable shocks that affect the network as a whole. Such an approach proves also useful to tackle one further empirical concern stemming from the fact that each agent's peer group (rather than the whole network) might be affected by common unobservable factors. Indeed, once our particularly large information on individual (observed) variables and network characteristics are taken into account, (within network) linking decisions appear uncorrelated with peer group-level observables. Finally, the variety of questions in the AddHealth questionnaire allows us to find observable proxies for typically unobserved individual characteristics that are commonly believed to induce self-selection (ability, leadership propensity, parental care etc.). The addition of school dummies is used to control for school-specific inputs.

Observe that school-dummies also account for differences in the strictness of anti-crime regulations across schools as well as for local crime policies. The identification of deterrence effects on crime is a difficult empirical exercise because of the well-known potential simultaneity and reverse causality issues (Levitt, 1997), which cannot totally be solved using our network-based approach. We avoid to directly estimate such effects (i.e. to include in the model specification observable measures of deterrence, such as local police expenditures or the arrest rate in the local area). Rather, we focus our attention on the estimation of peer effects on crime, once deterrence effects have been controlled for.

This strategy leads to the following main findings: conformity plays an important role in explaining criminal behavior of adolescents, especially for petty crimes. Specifically, a one-standard deviation increase in individual i's taste for conformity or equivalently in the average criminal activity of individual i's reference group raises individual i's level of crime by about 5.2 percent of a standard deviation when total crime is considered. It ranges from 9.8 to 1.4 moving from petty crimes to more serious crimes.

The rest of the paper can be described as followed. In the next section, we derive our main theoretical results. Section 3 describes the data and the empirical strategy. In section 4, we present our empirical results, both for all crimes and for each type of crime. Finally, section 5 concludes.

2 Theory

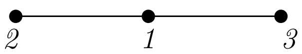
2.1 The basic model

There are N individuals/criminals in the economy.

The network $N = \{1, \ldots, n\}$ is a finite set of agents. The $n-$square adjacency matrix $\mathbf{G}$ of a network $\mathbf{g}$ keeps track of the direct connections in this network. Here, two players i and j are directly connected (i.e. best friends) in $\mathbf{g}$ if and only if $g_{ij} = 1$, and $g_{ij} = 0$, otherwise. Given that friendship is a reciprocal relationship, we set $g_{ij} = g_{ji}$. We also set $g_{ii} = 0$. The set of individual i's best friends (direct connections) is: $N_i(\mathbf{g}) = \{j \neq i \mid g_{ij} = 1\}$, which is of size g_i (i.e. $g_i = \sum_{j=1}^{n} g_{ij}$ is the number of direct links of individual i). This means in particular that, if i and j are best friends, then in general $N_i(\mathbf{g}) \neq N_j(\mathbf{g})$ unless the graph/network is complete (i.e. each individual is friend with everybody in the network). This also implies that groups of friends may overlap if individuals have common best friends. To summarize, the *reference group* of each individual i is $N_i(\mathbf{g})$, i.e. the set of his/her best friends, which does not include him/herself.

Let $\gamma_{ij} = g_{ij}/g_i$, for $i \neq j$, and set $\gamma_{ii} = 0$. By construction, $0 \leq \gamma_{ij} \leq 1$. Note that $\boldsymbol{\gamma}$ is a row-normalization of the initial friendship network $\mathbf{g}$, as illustrated in the following example, where $\mathbf{G}$ and $\boldsymbol{\Gamma}$ are the adjacency matrices of, respectively, $\mathbf{g}$ and $\boldsymbol{\gamma}$.

Example 1 *Consider the following friendship network* $\mathbf{g}$*:*



Then,

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \textit{and} \quad \boldsymbol{\Gamma} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Preferences We focus on adolescent crime and we denote by $e_i(\mathbf{g})$ the crime effort level of criminal i in network $\mathbf{g}$. We also denote by $\overline{e}_i(\mathbf{g})$ the average crime effort of the g_i best friends of i, which is given by:

$$\overline{e}_i(\mathbf{g}) = \frac{1}{g_i} \sum_{j=1}^{j=n} g_{ij} e_j \tag{1}$$

From now on, when there is no risk of confusion, we drop the argument $\mathbf{g}$. Each individual/criminal selects an effort $e_i \geq 0$ and obtains a payoff $u(e_i, \overline{e}_i)$ given by the following utility function[7]:

$$u_i(e_i, \overline{e}_i) = a + b_i\, e_i - p\, e_i\, f - c\, e_i^2 - d\, (e_i - \overline{e}_i)^2 \tag{2}$$

with $a, c, d > 0$, and $b_i > 0$ for all i.

This utility has a standard cost/benefit structure (as in Becker, 1968). The proceeds from crime is given by $a + b_i\, e_i$ and is increasing in own effort e_i. There is an ex ante *idiosyncratic heterogeneity*, b_i, which captures the fact that individuals differ in their ability (or productivity) of committing crime. Indeed, for a given effort level e_i, the higher b_i, the higher the productivity and thus the higher the booty $a + b_i\, e_i$. Observe that b_i is correlated with individual i's idiosyncratic characteristics such as his/her parents' education, the neighborhood where he/she lives, age, sex, race, etc., but also with the average characteristics of individual i's best friends, i.e. average level of parental education of i's friends, etc. (contextual effects). To be more precise, b_i can be written as:

$$b_i(\mathbf{x}) = \sum_{m=1}^{M} \beta_m x_i^m + \frac{1}{g_i} \sum_{m=1}^{M} \sum_{j=1}^{n} \theta_m g_{ij}\, x_j^m \tag{3}$$

where x_i^m is a set of M variables accounting for observable differences in individual, neighborhood and school characteristics of individual i, and β_m, θ_m are parameters.

The costs of committing crime are captured by the probability to be caught $p\, e_i$, which increases with own effort e_i, as the apprehension probability increases with one's involvement in crime, times the fine f, i.e. the severity of the punishment. Also, as it now quite standard (see e.g. Verdier and Zenou, 2004; Conley and Wang, 2006), individuals have a *moral* cost of committing crime equal to $c\, e_i^2$, which is reflected here by their degree of honesty c.[8] So the higher c, the higher the moral cost and it increases with crime effort.

Finally, the new element in this utility function is the last term $d\, (e_i - \overline{e}_i)^2$, which reflects the influence of friends' behavior on own action. It is such that each individual wants to minimize the social distance between him/herself and his/her reference group, where d is the parameter describing the taste for conformity. Indeed, the individual loses utility $d(e_i - \overline{e}_i)^2$ from failing to conform to others. This is the standard way economists have been modelling conformity (see, among others, Akerlof, 1980, Ballester et al., 2006, Bernheim, 1994, Kandel

[7] Our approach differs from that of Ballester et al. (2006) in that we consider both the $\mathbf{G}$ and $\mathbf{\Gamma}$ adjacency matrices while they deal only with the $\mathbf{G}$ matrix. It is because we focus on conformity, i.e. distance from the mean effort , while they mainly analyze the impact of peer effects, i.e. sum of efforts.

[8] Assuming different degrees of honesty c_i does not change our results.

and Lazear, 1992, Akerlof, 1997, Fershtman and Weiss, 1998). We can analyze the bilateral influences of this utility function. They are given by:

$$\frac{\partial^2 u_i(e_i, \overline{e}_i)}{\partial e_i \partial e_j} = \begin{cases} -2\,(c+d) < 0, \text{ when } i = j \\ 0, \text{ when } i \neq j \text{ and } g_{ij} = 0 \\ 2d > 0, \text{ when } i \neq j \text{ and } g_{ij} = 1 \end{cases} .$$

Since, when $i \neq j$, $2d > 0$, an increase in effort from j triggers a downwards shift in i's response and thus efforts are strategic complements from i's perspective within the pair (i, j).

Observe that beyond the idiosyncratic heterogeneity, b_i, there is a second type of heterogeneity, referred to as *peer heterogeneity*, which captures the differences between individuals due to network effects. Here it means that individuals have different types of friends and thus different reference groups $\overline{e}_i$. As a result, the social norm each individual i faces is endogenous and depends on his/her location in the network as well as the structure of the network. Indeed, in a star-shaped network (as the one described in Figure 1) where each individual is at most distance 2 from each other, the value of the social norm will be very different than a circle network, where the distance between individuals can be very large.

2.2 A simple symmetric case

In this section, we assume that, ex ante, all individuals/criminals are identical, i.e. same ex ante idiosyncratic heterogeneity, so that $b_i = b$.[9] This of course does not mean that they have the same peer heterogeneity since individuals have different reference groups.

We can calculate the Nash equilibrium of this game where each individual chooses e_i by taking as given the actions of the other players. We have the following result:

Proposition 1 *Assume that $b_i = b$ and $b > p\,f$. Then, the conformity game with payoffs (2) has a unique Nash equilibrium in pure strategies, which is given by:*

$$e_i^* = \overline{e}_i^* = \frac{b - p\,f}{2c} \tag{4}$$

In particular, the higher the deterrence, the lower the crime level.

Proof. See Appendix 1.

This is an interesting and surprising result. It says that, even if individuals are ex ante heterogeneous because of their location in the network and thus have different reference

[9] We relax these assumptions in the next section.

groups and social norms (peer heterogeneity), in a conformist equilibrium where each individual would like to conform as much as possible to the norm of his/her reference group, all individuals will exert the same effort level. This effort is increasing in the booty b, decreasing in the deterrence $p\,f$ and in the disutility of committing crime c. In other words, ex ante heterogeneity and the distribution (in particular the variance) of population do not matter in a conformist equilibrium even if it does ex ante. It is really the average that plays a crucial role in this model. This contrasts with the results of Ballester et al. (2006) who find that, when the utility function has not this conformism component, ex ante heterogenous agents are ex post heterogenous in terms of outcomes. So basically, the present mode shows that what is crucial to understand crime behavior among teenagers is their reference group and thus the friends they associate with. The level of criminal activities of the group will then be determined by the level of deterrence $p\,f$.

2.3 The general model

Let us generalize this theoretical model for the case of ex ante heterogenous individuals in terms of b_i. We have the following result:

Proposition 2 *Consider the general case when all individuals have ex ante idiosyncratic and peer heterogeneities, and different tastes for conformity. Assume that $b_i > p\,f$ for all i. Then, there exists a unique Nash equilibrium where each individual i provides the following crime effort:*

$$\begin{aligned} e_i^* &= \frac{d}{c+d}\,\overline{e}_i + \frac{b_i}{2\,(c+d)} - \frac{p\,f}{2\,(c+d)}, \\ &= \left(\frac{d}{c+d}\right)\frac{1}{g_i}\sum_{j=1}^{j=n} g_{ij}\,e_j + \frac{b_i}{2\,(c+d)} - \frac{p\,f}{2\,(c+d)}, \end{aligned} \tag{5}$$

which is increasing with the average crime effort of the reference group $\overline{e}_i$ and decreasing with deterrence pf, i.e.

$$\frac{\partial e_i^*}{\partial \overline{e}_i} > 0 \text{ and } \frac{\partial e_i^*}{\partial pf} < 0 \tag{6}$$

Proof. See Appendix 1.

The previous result of Proposition 1 does not hold anymore since there are now both *idiosyncratic* and *peer heterogeneities.* We obtain that, as long as peer effects matter, individuals will provide criminal effort proportional to their reference group and to their ex ante idiosyncratic heterogeneity b_i. Also, deterrence $p\,f$ will negatively affect the crime effort.

Thus, not surprisingly, Proposition 2 shows that the only Nash equilibrium is asymmetric since each individual provides different crime efforts. Also, the higher their taste for conformity d, the higher is this effect since peer effects are even more important. Here, people conform to the norm because it is costly for them to deviate from it but the norm (i.e. the crime level of the reference group) is negatively affected by deterrence. Interestingly, in (5), we are able to decompose the equilibrium crime effort e_i^* into an ex ante idiosyncratic heterogeneity effect, a deterrence effect and a peer effect.

3 Data and empirical strategy

3.1 Data

Our data source is the National Longitudinal Survey of Adolescent Health (AddHealth), which contains detailed information on a nationally representative sample of 90,118 students in roughly 130 private and public schools, entering grades 7-12 in the 1994-1995 school year.[10] AddHealth contains unique information on friendship relationships, which is crucial for our analysis. The friendship information is based upon actual friends' nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[11] A link exists between two friends if at least one of the two individuals has identified the other as his/her best friend.[12]

Figure 1 shows the empirical distribution of friendship networks in our sample by their size (i.e. the number of network members).[13] It appears that most friendship networks have between 36 and 74 members The minimum number of friends in a network is 18, while the maximum is 88. The average and the standard deviation of network size are 49.51 and 16.80.

[10] For a detailed description of the survey and data, see the AddHealth website at: http://www.cpc.unc.edu/projects/addhealth.

[11] The limit in the number of nominations is not binding, not even by gender. Less than 1 percent of the students in our sample show a list of ten best friends, less than 3 percent a list of five males and roughly 4 percent name five females. On average, they declare to have 6.04 friends with a small dispersion around this mean value (standard deviation equal to 1.32).

[12] Note that, when an individual i identifies a best friend j who does not belong to the surveyed schools, the database does not include j in the network of i; it provides no information about j. Fortunately, in the large majority of cases (more than 93%), best friends tend to be in the same school and thus are systematically included in the network.

[13] The histograms show on the horizontal axes the percentiles of the empirical distribution of network component members corresponding to the percentages 1, 5, 10, 25, 50, 75, 90, 95, 100 and in the vertical axes the number of networks having number of members between the i and $i-1$ percentile.

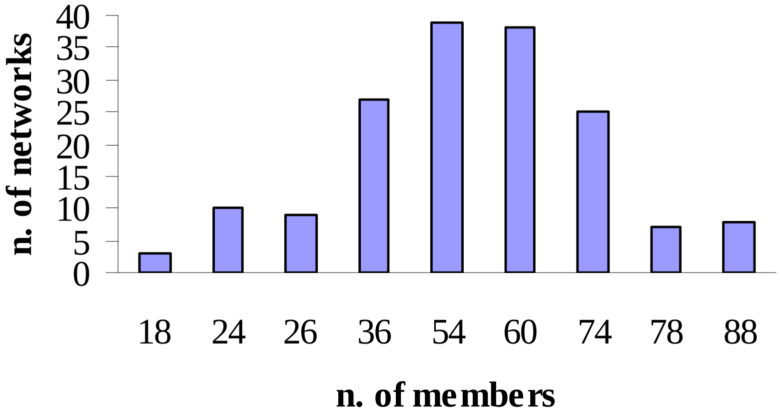


Figure 1. The empirical distribution of adolescent networks

By matching the identification numbers of the friendship nominations to respondents' identification numbers, one can obtain information on the characteristics of nominated friends.

Besides information on family background, school quality and area of residence, the AddHealth contains sensitive data on sexual behavior (contraception, pregnancy, AIDS risk perception), tobacco, alcohol, drugs and crime of a subset of adolescents. We use these data to construct our dependent variable e_i. Addhealth contains an extensive set of questions on juvenile delinquency, ranging from light offenses that only signal the propensity towards a delinquent behavior to serious property and violent crime.[14] Firstly, we adopt the standard approach in the sociological literature to derive an index of delinquency involvement based on self-reported adolescents' responses to a set of questions describing participation in a series of criminal activities. The survey asks students how often they participate in each of

[14] Specifically, it contains information on 15 delinquency items. Namely, paint graffiti or signs on someone else's property or in a public place; deliberately damage property that didn't belong to you; lie to your parents or guardians about where you had been or whom you were with; take something from a store without paying for it; get into a serious physical fight; hurt someone badly enough to need bandages or care from a doctor or nurse; run away from home; drive a car without its owner's permission; steal something worth more than $50; go into a house or building to steal something; use or threaten to use a weapon to get something from someone; sell marijuana or other drugs; steal something worth less than $50; take part in a fight where a group of your friends was against another group; act loud, rowdy, or unruly in a public place.

the different activities during the past year.[15] Each response is coded using an ordinal scale ranging from 0 (i.e. never participate) to 1 (i.e. participate 1 or 2 times), 2 (participate 3 or 4 times) up to 3 (i.e. participate 5 or more times). On the basis on these variables, a composite score is calculated for each respondent.[16] The mean is 1.03, with considerable variation around this value (the standard deviation is equal to 1.22). The Crombach$-\alpha$ measure is then used to assess the quality of the derived index. In our case, we obtain an α equal to 0.76 ($0 \leq \alpha \leq 1$) indicating that the different items incorporated in the index have considerable internal consistency. Secondly, in Section 4.2. we consider different categories of crime, which are chosen accordingly to the seriousness of the crime committed. Using the correspoding information for nominated friends, we are able, for each individual i, to calculate the average crime effort $\overline{e}_i$ of his/her peer group. Excluding the individuals with missing or inadequate information, we obtain a final sample of 9,322 students distributed over 166 networks.[17]

3.2 Empirical strategy

Guided by Proposition 2, our aim is to assess the actual empirical relationship between the group criminal effort $\overline{e}_i$ and individual effort level e_i^* (comparative statics result (6)).

The main novel feature of our estimation with respect to previous works is the use of the architecture of networks to evaluate peer effects. Let us explain this more clearly.

Reflection problem Without networks, it is difficult to differentiate between the effect of peers' choice of effort and peers' characteristics that do impact on their effort choice (the so-called *reflection problem*; Manski, 1993). Basically, the reflection problem arises because, in the standard approach, individuals interact in groups, that is individuals are affected by all individuals belonging to their group and by nobody outside the group. In other words, groups do not overlap. In the case of networks, as in our theoretical model, this is nearly never true since the reference group is the number of friends each individual has. So for example take individuals i and k such that $g_{ik} = 1$. Then, individual i is directly influenced by $g_i = \sum_{j=1}^{n_i} g_{ij} e_j$ while individual k is directly influenced by $g_k = \sum_{j=1}^{n_k} g_{kj} e_j$, and there is little chance for these two values to be the same unless the network is complete (i.e. everybody

[15] Respondents listened to pre-recorded questions through earphones and then they entered their answers directly on laptop computers. This administration of the survey for sensitive topics minimizes the potential for interview and parental influence, while maintaining data security.

[16] This is a standard factor analysis, where the factor loadings of the different variables are used to derive the total score.

[17] The networks include both criminals and noncriminals.

is linked with everybody). Formally, social effects are identified (i.e. no reflection problem) if $\mathbf{G}^2 \neq \mathbf{0}$, where $\mathbf{G}^2$ keeps track of indirect connections of length 2 in $\mathbf{g}$.[18] This condition guarantees that $\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^2$ are linearly independent. $\mathbf{G}^2 \neq \mathbf{0}$ means that there exist at least a path of length 2 between two individuals.[19] In other words, if i and j are friends and j and k are friends, it does not necessarily imply that i and k are also friends. The reason for identification is now straightforward. Since individual $k \notin g_i$, the characteristics of k do not directly affect e_i (i's outcome) but, since $k \in g_j$, they affect e_j (j's outcome), and since $j \in g_i$, e_j affects e_i. As a result, the characteristics of k affects e_i only *indirectly* through its effect on e_j. This means that the characteristics of k are a valid instrument to estimate the endogenous social effect for e_i. These results are formally derived in Bramoullé et al. (2006) (see, in particular Proposition 3) and used in Calvó-Armengol et al. (2008). Lavescher (2005) is, to our knowledge, the only other empirical analysis using a multiple group strategy. Cohen-Cole (2006) presents a similar argument, i.e. the use of out-group effects, to achieve the identification of the endogenous group effect in the linear-in-means model.

Endogenous network formation/correlated effects Although this setting allows us to solve the reflection problem, the estimation results might still be flawed because of the presence of unobservable factors affecting both individual and peer behavior. It is thus difficult to disentangle the endogenous peer effects from the correlated effects, i.e. from effects arising from the fact that individuals in the same group tend to behave similarly because they face a common environment. If individuals are not randomly assigned into groups, this problem might originate from the possible sorting of agents. If the variables that drive this process of selection are not fully observable, potential correlations between (unobserved) group-specific factors and the target regressors are major sources of bias. In our case, two types of possibly correlated effects arise, i.e. at the network level and at the peer group level. The use of network fixed effects proves useful in this respect. Assume, indeed, that agents self-select into different networks in a first step, and that link formation takes place within networks in a second step. Then, as Bramoullé et al. (2006) observe, if linking decisions are uncorrelated with the observable variables, this two-step model of link formation generates network fixed effects. Assuming additively separable network heterogeneity, a within group specification is able to control for these correlated effects. In other words, we use a model

[18] The coefficient $g_{ij}^{[2]}$ in the (i,j) cell of $\mathbf{G}^2$ gives the number of paths of length 2 in $\mathbf{g}$ between i and j.

[19] It is extremely rare that in the real world the condition $\mathbf{G}^2 \neq \mathbf{0}$ is not satisfied since it would basically imply that all networks are complete. In our dataset, where 166 networks are considered (see above in the data section), none of them are complete and all satisfy the condition that guarantees the identification of social effects.

specification with a network-specific component of the error term, and adopt a traditional (pseudo) panel data fixed effects estimator, namely, we subtract from the individual-level variables the network average.[20]

Observe that our particularly large information on individual (observed) variables should reasonably explain the process of selection into groups. Then, the inclusion of network fixed effects acts as a further control for possible sorting effects based on unobservables.

To document to what extent this approach accounts for self-selection in our case, we need to provide evidence that (*i*) network-fixed effects account for unobservable factors driving the allocations of agents into networks and (*ii*) once observables and network-fixed effects are controlled for linking decisions are uncorrelated with peer-level observables.

We thus consider individual variables that are commonly believed to induce self-selection into teenagers' friendship groups and perform two different exercises. Firstly, we estimate the correlations between such individual-level variables and the network averages of the residuals obtained from a regression analysis where the influence of a variety of other factors (see Table A.1, Appendix 2 for a precise description of variables) and network-fixed effects are washed out. Secondly, we estimate the correlations between such individual-level variables and peer-group averages (i.e., averages over best friends), once the influence of our extensive set of controls and network-fixed effects are washed out. The results are reported in Table 1 (in the second and third column, respectively). The estimated correlation coefficients are not statistically significant for all attributes considered in both columns. This indicates that, in our case, (*i*) the particularly large information on individual (observed) variables and (additively separable) unobserved network characteristics account for a possible sorting of students into networks and (*ii*) conditionally on individual and network characteristics, linking decisions are uncorrelated with observable variables.

[*Insert Table* 1 *here*]

Correlated individual effects Finally, one might question the presence of problematic unobservable factors that are nor network-specific nor peer-group-specific, but rather individual-specific. In this respect, the richness of the information provided by the AddHealth questionnaire on adolescents' behavior allow us to find proxies for typically unobserved in-

[20] Bramoullé et al. (2006) also deal with this problem in the context of networks. In their Proposition 5, they show that if the matrices $\mathbf{I}$, $\mathbf{G}$, $\mathbf{G}^2$ and $\mathbf{G}^3$ are linearly independent, then by subtracting from the variables the network component average (or the average over neighbors, i.e. direct friends) social effects are again identified and one can disentangle endogenous effects from correlated effects. In our dataset this condition of linear independence is always satisfied.

dividual characteristics that may be correlated with our variable of interest. Specifically, to control for differences in leadership propensity across adolescents we include an indicator of self-esteem and an indicator of the level of physical development compared to the peers, and we use mathematics score as an indicator of ability. Also, we attempt to capture differences in attitude towards education and parenting by including indicators of the student's motivation in education and parental care.

Correlated school effects Similar arguments can be put forward for the existence of possible correlations between our variable of interest and unobservable school characteristics affecting structure and/or quality of school-friendship networks in analyzing students' school performance. Because the AddHealth survey interviews all children within a school, we estimate our model conditional on school fixed effects (i.e. we incorporate in the estimation *school dummies*). This approach enables us to capture the influence of school level inputs (such as teachers and students quality, and possibly the parents' residential choices), so that only the variation in the average behavior of peers (across students in the same school) would be exploited.[21]

Deterrence effects So far in this section, we have focused our attention on the main purpose of our empirical analysis, which is to be found in the identification of peer effects and conformism in crime using the network architecture. The identification of deterrence effects (pf in our theoretical model) on crime is an equally difficult empirical exercise because of the well-known potential simultaneity and reverse causality issues (Levitt, 1997), which cannot be totally solved using our network based empirical strategy. School dummies, however, also account for differences in the strictness of anti-crime regulations across schools (i.e. differences in the expected punishment for a student who is caught possessing illegal drug, stealing school property, verbally abusing a teacher, etc.) as well as for differences in crime policies at the local level (because schools are in different areas). As a result, instead of directly estimating deterrence effects (i.e. to include in the model specification observable measures of deterrence, such as local police expenditures or the arrest rate in the local area), we focus our attention on the estimation of peer effects in crime, accounting for observable and unobservable school, and hence area-of-residence, variables (such as policing practicing, ethnic concentration, low informal social control, lack of educational or economic opportunities,etc....) that might be correlated with our variable of interest.

Assuming n_κ individuals in each of the K networks in the economy, for $i = 1, ..., n_\kappa$,

[21] Most of the times (but not always) school dummies coincide with network dummies. The introduction of student-grade or student-year of attendance dummies does not change qualitatively the results on our target variable.

$\kappa = 1, ..., K$, and using (3), the econometric counterpart of (5) is given by:

$$e_{i,\kappa} = \phi \frac{1}{g_{i,\kappa}} \sum_{j=1}^{n_{i,\kappa}} g_{ij,\kappa}\, e_{j,\kappa} + \sum_{m=1}^{M} \beta_1^m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \theta_m g_{ij,\kappa} x_{j,\kappa}^m + \eta_k + \varepsilon_{ik} \qquad (7)$$

where $e_{i,k}$, is the index of criminality of individual i in network κ, $x_{i,\kappa}^m$ (for $m = 1, ..., M$) is the set of M control variables containing an extensive number of individual, family, school and residential area characteristics, $g_{i,\kappa} = \sum_{j=1}^{n_\kappa} g_{ij,\kappa}$ is the number of direct links of i, $\sum_{j=1}^{n_\kappa} \left(g_{ij,\kappa} x_{j,\kappa}^m\right) / g_{i,\kappa}$ is the set of the average values of the M controls of i's direct friends (i.e. contextual effects). As stated in the theoretical model, $\sum_{m=1}^{M} \beta_1^m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \gamma_m g_{ij,\kappa} x_{j,\kappa}^m$ reflects the ex-ante idiosyncratic heterogeneity of each individual i and our measure of *taste for conformity* or *strength of peer effects* is captured by the parameter ϕ (in the theoretical model $\phi = d/\left(c + d\right)$). The error term consists of a network specific component (constant over individuals in the same network), which might be correlated with the regressors, η_k, and a white noise component, ε_{ik}. A precise description of the variables included and the corresponding descriptive statistics are contained in the Data Appendix to this paper (Table A.1, Appendix 2).

Model (7) is the empirical equivalent of the first order conditions of our model of network peer effects given by (5) in Proposition 2. It is the so-called *spatial lag model* or *mixed-regressive spatial autoregressive model* (Anselin, 1988) with the addition of a network-specific component of the error term. Once the variables are transformed in deviations from the network-specific means, a Maximum Likelihood approach (see, e.g. Anselin, 1988) allows us to estimate jointly $\widehat{\boldsymbol{\beta}}$, $\widehat{\boldsymbol{\gamma}}$, and $\widehat{\phi}$.

4 Empirical results

Testing the model The maximum likelihood estimation results of model (7) are reported in the second column of Table 2 ("All crimes").[22] The table shows that the estimated coefficient of ϕ, which measures the *taste for conformity*, is statistically significant and has a positive sign. Specifically, a one-standard deviation increase in individual i's taste for conformity or equivalently in the average criminal activity of individual i's reference group raises individual i's level of crime by about 5.2 percent of a standard deviation. This evidence supports our theoretical framework predicting a relevant role of peers and conformity to peers' behavior

[22] The qualitative results on our target variable are robust to alternative sets of control variables. The complete lists of estimation results are available upon request.

in shaping criminal activities among teenagers.

[*Insert Table* 2 *here*]

Different types of crime The literature on local interactions has uncovered some interesting differences between different types of crime. For instance, Ludwig et al. (2000) find that neighborhood effects are large and negative for violent crime but have a mild positive effect on property crime. In contrast, Glaeser et al. (1996) find instead that social interactions seem to have a large effect on petty crime, a moderate effect on more serious crime and a negligible effect on very violent crime.

The basic idea of our theoretical model is that agents' criminal behavior is driven by their desire to reduce the discrepancy between their own crime effort and that of their reference group (i.e. their best friends). We find that such a model is validated by our data for juvenile crime as a whole.

The richness of the information provided by the AddHealth data on juvenile crime enables us also to test our conformism model for different types of crime, thus making our analysis directly comparable to previous works. Specifically, we analyze whether the magnitude of the peer effects depends on the type of crime committed. We split the offences reported in our data in three groups (with increasing costs of committing crime). The first group (*type-1 crimes*) contains (*i*) to paint graffiti or sign on someone else's property or in a public place; (*ii*) to lie to the parents or guardians about where or with whom having been; (*iii*) to run away from home; (*iv*) to act loud, rowdy, or unruly in a public place. The second group (*type-2 crimes*) consists of (*i*) to get into a serious physical fight; (*ii*) to hurt someone badly enough to need bandages or care from a doctor or nurse; (*iii*) to drive a car without its owner's permission; (*iv*) to steal something worth less than $50. The third group (*type-3 crimes*) encompasses (*i*) to take something from a store without paying for it; (*ii*) to steal something worth more than $50; (*iii*) to go into a house or building to steal something; (*iv*) to use or threat to use a weapon to get something from someone; (*v*) to sell marijuana or other drugs. Less than 20 percent of the teenagers in our sample confess to have committed the more serious offences.[23] To be precise, these three groups contain 3,488, 4,084 and 1,803 individuals respectively.

We estimate the following modified version of model (7):

$$e_{i,\kappa,l} = \alpha\phi_l \sum_{j=1}^{n_{i,\kappa}} g_{ij,\kappa}\, e_{j,\kappa} + \sum_{m=1}^{M} \beta^m x_{i,\kappa}^m + \frac{1}{g_{i,\kappa}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \theta_m g_{ij,\kappa} x_{j,\kappa}^m + \eta_k + \varepsilon_{i,k,l}$$

[23] Adolescents are selected in a more serious type of crime group if they have committed at least one of the offences considered in the group.

where $e_{i,k,l}$ is now the index of crime of type l committed by individual i in network κ, and the rest of the notation defined for model (7) applies. The estimation of this model provides type of crime-specific peer effects. The results are contained in columns three, four and five of Table 2. We find that the estimated coefficient ϕ_l, which measures the taste for conformity for type-l crime, remains always significant and positive whatever the seriousness of the crime considered, but it decreases in magnitude when moving from light to more serious crimes. A one-standard deviation increase in individual i's taste for conformity for type-1 crimes or equivalently a one-standard deviation increase in the average criminal activity of individual i's reference group translates roughly into a 9.8 percent decrease in standard deviations of individual i's criminal activity when petty crimes (*type-1 crimes)* are considered, whereas this effect amounts to 6.3 and only to 1.4 for intermediary (*type-2 crimes)* and serious crimes (*type-3 crimes),* respectively. This evidence is in line with the findings of Glaeser et al. (1996) who show that social interactions are more important for petty crimes.

5 Conclusion and policy implications

In education, crime, smoking, teenage pregnancy, school dropout, etc. economists have pointed out the importance of peer effects in explaining these outcomes (see e.g. Glaeser and Scheinkman, 2001). To understand the generating mechanism of such peer effects is essential for the interpretation of the findings and thus to provide policy guidance. Conformity is key to determine economic outcomes that involve interactions with peers. In the present paper, we propose a model that explains how conformity and deterrence impact on criminal activities. In particular, we find significant impact of peers on individual criminal activity for individuals belonging to the same group of friends. We then test the model using the U.S. National Longitudinal Survey of Adolescent Health (AddHealth), which contains unique detailed information on friendship relationships among delinquent teenagers. A "reversion-to-the-group-mean" effect is identified. We find that conformity is very strong within groups of friends, especially for petty crimes.

Our results suggest that, for teenagers, the decision to commit crime depends on the seriousness of crime. In particular, for petty crimes, adolescents are strongly affected by their environment and peers because of externalities involved in social-decision making. In their study of a gang located in a black inner-city neighborhood, Levitt and Venkatesh (2000) find that "social/nonpecuniary factors are likely to play an important role" in criminal decisions and gang activities. Here, even though we do not focus on gangs, we highlight one of these social/nonpecuniary factors: the desire to conform to the group's norm. Because of the implications of juvenile crime for adolescent's behavior in the future, an effective policy should not only be measured by the possible crime reduction it implies but also by the group interactions it engenders.

To be more precise, if social interactions and conformism are crucial to understand juvenile criminal activities, then a targeted policy identifying "key players" in a given area (Ballester et al., 2006, 2008) may be an effective way to reduce crime. A key player (or a key group) is an individual (or a group of persons) belonging to a network of criminals who, once removed, leads to the highest aggregate delinquency reduction. In practice, the planner may want to identify optimal network targets to concentrate (scarce) investigatory resources on some particular individuals, or to isolate them from the rest of the group, either through leniency programs, social assistance programs, or incarceration. The success of such policy depends on the ability to identify a social network and this task may be not as difficult as it seems to be. For instance, Haynie (2001) and the present paper use friendship data to identify delinquent peer networks for adolescents in the U.S. that participated in an in-school survey in the 1990's. Sarnecki (2001) provides a comprehensive study of co-offending relations and corresponding network structure for football hooligans and right-wing extremists in Stockholm. In all these cases, one may directly use the available data to determine the key player or group players.

References

[1] Akerlof, G.A. (1980), "A theory of social custom of which unemployment may be one consequence," *Quarterly Journal of Economics* 94, 749-775.

[2] Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 1005-1027.

[3] Anselin, L. (1988), *Spatial Econometrics, Methods and Models*, Dordrech: Kluwer Academics.

[4] Anselin, L. (1995), *Spacestat Version 1.80 User's Guide. Regional Research Institute*, West Virginia University, Morgantown WV.

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

[6] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2008), "Delinquent networks," Unpublished manuscript, Stockholm University.

[7] Battu, H., Mwale, M. and Y. Zenou (2007), "Oppositional identities and the labor market," *Journal of Population Economics* 20, 643-667.

[8] Bayer, P., Hjalmarsson, R. and D. Pozen (2007), "Building criminal capital behind bars: Peer effects in juvenile corrections," NBER Working Paper No. 12932.

[9] Becker, G. (1968), "Crime and punishment: An economic approach," *Journal of Political Economy* 76 169-217.

[10] Bernheim, B.D. (1994), "A theory of conformity," *Journal of Political Economy* 102, 841-877.

[11] Bramoullé, Y., Djebbari, H. and B. Fortin (2006), "Indentification of peer effects through social networks," Unpublished manuscript, Université Laval.

[12] Burdett, K., Lagos, R. and R. Wright (2003), "Crime, inequality, and unemployment," *American Economic Review* 93, 1764-1777.

[13] Burdett, K., Lagos, R. and R. Wright (2004), "An on-the-job search model of crime, inequality, and unemployment," *International Economic Review* 45, 681-706.

[14] Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2008), "Peer effects and social networks in education," Unpublished manuscript, Stockholm University.

[15] Calvó-Armengol, A., Verdier, T. and Y. Zenou (2007), "Strong and weak ties in employment and crime," *Journal of Public Economics* 91, 203-233.

[16] 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, 935-954.

[17] Case, A.C. and L.F. Katz (1991), "The company you keep: the effects of family and neighborhood on disadvantaged youths," NBER Working Paper No. 3705.

[18] Cohen-Cole, E. (2006), "Multiple groups identification in the linear-in-means model," *Economics Letters* 92, 157-162.

[19] Conley, J.P. and P. Wang (2006), "Crime and ethics," *Journal of Urban Economics* 60, 107-123.

[20] Durlauf, S.E. (2004), "Neighborhood effects," In: *Handbook of Regional and Urban Economics Vol. 4*, J.V. Henderson and J-F. Thisse (Eds.), Amsterdam: Elsevier Science, pp. 2173-2242.

[21] Durlauf, S.N. and R.A. Moffitt (2003), "Special issue on empirical analysis of social interactions," *Journal of Applied Econometrics* 18, 499-619.

[22] Ferrer, R. (2008), "Breaking the law when others do: A model of law enforcement with neighborhood externalities," Unpublished manuscript, Vanderbilt University.

[23] Fershtman, C. and Y. Weiss (1998), "Social rewards, externalities and stable preferences," *Journal of Public Economics* 70, 53-73.

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

[25] Glaeser, E.L. and J. Scheinkman (2001), "Measuring social interactions," In: *Social Dynamics*, S.N. Durlauf and H.P. Young (Eds.), Cambridge: MIT Press, pp. 83-132.

[26] Haynie, D.L. (2001) "Delinquent peers revisited: Does network structure matter?," *American Journal of Sociology* 106, 1013-1057.

[27] Huang, C-C., Laing, D. and P. Wang (2004), "Crime and poverty: A search-theoric approach," *International Economic Review* 45, 909-938.

[28] İmrohoroğlu, A., Merlo, A. and P. Rupert (2000), "On the political economy of income redistribution and crime," *International Economic Review* 41, 1-25.

[29] Jackson, M. O. (2007), "The economics of social networks," In: *Proceedings of the Ninth World Congress of the Econometric Society Vol. 1*, R. Blundell, W.K. Newey and T. Persson (Eds.), Cambridge: Cambridge University Press, pp. 1-56.

[30] Jacob, B.A. and L. Lefgren (2003), "Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime," *American Economic Review* 93, 1560-1577.

[31] Kandel, E. and E.P. Lazear (1992), "Peer pressure and partnerships," *Journal of Political Economy* 100, 801-817.

[32] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood effects on crime for female and male youth: evidence from a randomized housing voucher experiment," *Quarterly Journal of Economics* 120, 87-130.

[33] Laschever, R. (2005), "The doughboys networks: Social interactions and labor market outcomes of World War I veterans," Unpublished manuscript, Cornell University.

[34] Levitt, S.D. (1997), "Using electoral cycles in police hiring to estimate the effect of police on crime," *American Economic Review* 87, 270-90.

[35] Levitt, S.D. (1998), "Juvenile crime and punishment," *Journal of Political Economy* 106, 1156-1185.

[36] Levitt, S.D. and L. Lochner (2001), "The determinants of juvenile crime," In: *Risky Behavior among Youths*, J. Gruber (Ed.), Chicago: Chicago University Press, pp. 327-374.

[37] Levitt, S.D. and S.A. Venkatesh (2000), "An economic analysis of a drug-selling gang's finances," *Quarterly Journal of Economics* 115, 755-789.

[38] Lochner, L. (2004), "Education, work, and crime: A human capital approach," *International Economic Review* 45, 811-843.

[39] Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001), "Urban poverty and juvenile crime: evidence from a randomized housing-mobility experiment," *Quarterly Journal of Economics* 116, 655-679.

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

[41] Matsueda, R.L. and K. Anderson (1998), "The dynamics of delinquent peers and delinquent behavior," *Criminology* 36, 269-308.

[42] Mocan, H.N. and D.I. Rees (2005), "Economic conditions, deterrence and juvenile crime: Evidence from micro data," *American Law and Economics Review* 7, 319-349.

[43] Patacchini, E. and Y. Zenou (2008), "The strength of weak ties in crime," *European Economic Review* 52, 209-236.

[44] Sah, R. (1991), "Social osmosis and patterns of crime," *Journal of Political Economy*, 99, 1272-1295.

[45] Sarnecki, J. (2001), *Delinquent Networks: Youth Co-Offending in Stockholm*, Cambridge: Cambridge University Press.

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

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

[48] Warr, M. (1996), "Organization and instigation in delinquent groups," *Criminology* 34, 11-37.

[49] Warr, M. (2002), *Companions in Crime: The Social Aspects of Criminal Conduct*, Cambridge: Cambridge University Press.

Appendix 1: Proofs of propositions of the model

Proof of Proposition 1. First, observe that $\mathbf{\Gamma}$ is a stochastic matrix, that is $\gamma_{ij}^{[k]} \geq 0$ and $\sum_j \gamma_{ij}^{[k]} = 1$, and thus the largest egenvalue of $\mathbf{\Gamma}$ is 1, i.e. $\mu_1(\mathbf{\Gamma}) = 1$. Second, by plugging (1) in (2) for the case $b_i = b$, we obtain:

$$\begin{aligned} u_i(e_i, \overline{e}_i) &= a + b\, e_i - p\, e_i\, f - c\, e_i^2 - d\,(e_i - \overline{e}_i)^2 \\ &= a + b\, e_i - p\, e_i\, f - c\, e_i^2 - d \left[e_i - \sum_{j=1}^{j=n_i(\mathbf{g})} \gamma_{ij} e_j \right]^2 \\ &= a - d \left[\sum_{j=1}^{j=n_i(\mathbf{g})} \gamma_{ij} e_j \right]^2 + (b - p\, f)\, e_i - (c+d)\, e_i^2 + 2d \sum_{j=1}^{j=n_i(\mathbf{g})} \gamma_{ij} e_i e_j \end{aligned}$$

Now, assuming $b > p\, f$, we can apply Theorem 1 of Ballester et al. (2006)[24] with $\alpha = b - p\, f$, $\beta = 2\,(c+d)$, $\gamma = 0$[25] and $\lambda = 2d$. Hence, the condition for existence and uniqueness of a Nash equilibrium can be written as: $2(c+d) > 2d$, which is always satisfied since $c > 0$. Third, let us calculate the Bonacich vector. By definition,

$$\begin{aligned} b_i(\boldsymbol{\gamma}, \phi) &= m_{ii}(\boldsymbol{\gamma}, \phi) + \sum_{j \neq i} m_{ij}(\boldsymbol{\gamma}, \phi) \\ &= \phi \sum_{j=1}^{n} \gamma_{ij} + ... + \phi^k \sum_{j=1}^{n} \gamma_{ij}^{[k]} + ... \\ &= \sum_{j=1}^{+\infty} \phi^k \end{aligned}$$

since $\mathbf{\Gamma}$, $\mathbf{\Gamma}^1$,..., $\mathbf{\Gamma}^k$, ... are stochastic matrices and thus $\sum_{j=1}^{n} \gamma_{ij} = ... = \sum_{j=1}^{n} \gamma_{ij}^{[k]} = 1$. As a result,

$$b_i(\boldsymbol{\gamma}, \phi) = \sum_{j=1}^{+\infty} \phi^k = \frac{1}{1-\phi}$$

Applying again Theorem 1 in Ballester et al. (2006), where $\phi = d/(c+d)$, our Nash equilibrium is given by:

$$\mathbf{e}^* = \begin{pmatrix} \frac{b - p\, f}{2c} \\ ... \\ \frac{b - p\, f}{2c} \end{pmatrix}$$

[24] Observe that the term $a - d \left[\sum_{j=1}^{j=n_i(\mathbf{g})} \gamma_{ij} e_j \right]^2$ does not matter since the derivative of this term with respect to e_i is equal to zero.

[25] This is the γ in Ballester et al. (2006).

This implies that $\mathbf{e}^* = \overline{\mathbf{e}}^*$ and thus all players provide the same effort level $(b - pf) / (2c)$. ▮

Proof of Proposition 2. First, observe that $\mathbf{\Gamma}$ is a stochastic matrix ($\gamma_{ij} \geq 0$ and $\sum_j \gamma_{ij} = 1$) and thus its largest egenvalue is 1, i.e. $\mu_1(\mathbf{\Gamma}) = 1$. Second, as for the proof of Proposition 1, we have:

$$\begin{aligned} u_i(e_i, \overline{e}_i) &= a + b_i\, e_i - p\, e_i\, f - c\, e_i^2 - d_i\, (e_i - \overline{e}_i)^2 \\ &= a - d_i \left[\sum_{j=1}^{j=n} \gamma_{ij} e_j \right]^2 + (b_i - p\, f)\, e_i - (c + d_i)\, e_i^2 + 2d_i \sum_{j=1}^{j=n} \gamma_{ij} e_i e_j \end{aligned}$$

Assume that $b_i > p\, f$ for all i. The utility function is nearly the same as the one in Ballester et al. (2006)[26] where $\alpha_i = b_i - p\, f$, $\beta = 2\,(c + d)$, $\gamma = 0$[27] and $\lambda = 2d_i$. The main difference is that we now have ex ante heterogeneity because of α_i. However, because $\gamma = 0$ (i.e. there is no global substituability), the condition for existence and uniqueness of a Nash equilibrium is still given by $\beta > \gamma \mu_1(\mathbf{\Gamma})$,[28] which in our case is equivalent to: $2(c_i + d) > 2d$ for each i. This is always satisfied since $c_i > 0$ for all i. Third, (5) is just the first order condition for each individual i. ▮

[26] Observe that the term $a - d \left[\sum_{j=1}^{j=n_i(\mathbf{g})} \gamma_{ij} e_j \right]^2$ does not matter since the derivative of this term with respect to e_i is equal to zero.

[27] This is the γ in Ballester et al. (2006).

[28] See Ballester and Calvó-Armengol (2006).

Appendix 2: Data appendix

Table A.1: Description of Data (9,322 individuals, 166 networks)

	Variable definition	Mean	St.dev	Min	Max
Delinquency index	In the text	1.03	1.22	0	3
Delinquency index of best friends	Average value of the delinquency index over direct friends.	1.01	1.04	0	3
Delinquency index (type-1 crime)	In the text	1.66	1.15	0	3
Delinquency index (type-2 crime)	In the text	0.98	0.78	0	3
Delinquency index (type-3 crime)	In the text	0.59	0.33	0	3
Individual socio-demographic variables					
Female	Dummy variable taking value one if the respondent is female.	0.40	0.34	0	1
Age	Respondents' age measured in years.	15.25	1.85	10	19
Health status	Response to the question "In the last month, how often did a health or emotional problem cause you to miss a day of school", coded as 0= never, 1=just a few times, 2= about once a week, 3= almost every day, 4= every day.	3.03	1.74	0	4
Religion practice	Response to the question: "In the past 12 months, how often did you attend religious services", coded as 0= not applicable, 1= never, 2= less than once a month, 3= once a month or more, but less than once a week, 4= once a week or more.	2.69	0.78	0	4
Black or African American	Race dummies. “White” is the reference group.	0.20	0.31	0	1
Other races	“	0.10	0.13	0	1
School attendance	Number or years the respondent has been a student at the school.	3.29	1.86	1	6
Student grade	Grade of student in the current year.	9.24	3.14	6	13
Mathematics score	Score in mathematics at the most recent grading period, coded as1= D or lower, 2= C, 3=B, 4=A.	1.94	1.31	1	4
organized social participation	Dummy taking value one if the respondent participates in any clubs, organizations, or teams at school in the school year.	0.65	0.20	0	1
Motivation in education	Dummy taking value one if the respondent reports to try very hard to do his/her school work well, coded as 1=I never try at all, 2=I don't try very hard, 3=I try hard enough, but not as hard as I could, 4=I try very hard to do my best.	2.24	0.88	1	4
Physical development	Response to the question: "How advanced is your physical development compared to other boys your age", coded as 1= I look younger than most, 2= I look younger than some, 3= I look about average, 4= I look older than some, 5= I look older than most	3.12	2.51	1	5
Self esteem	Response to the question: "Compared with other people your age, how intelligent are you", coded as 1= moderately below average, 2= slightly below average, 3= about average, 4= slightly above average, 5= moderately above average, 6= extremely above average.	3.93	1.37	1	6
Family background variables					
Household size	Number of people living in the household.	3.50	1.73	1	6
Two married parent family	Dummy taking value one if the respondent lives in a household with two parents (both biological and non biological) that are	0.42	0.57	0	1

	married.				
Single parent family	Dummy taking value one if the respondent lives in a household with only one parents (both biological and non biological).	0.22	0.43	0	1
Public assistance:	Dummy taking value one if either the father or the mother receives public assistance, such as welfare.	0.12	0.16	0	1
Mother working	Dummy taking value one if the mother works for pay.	0.64	0.45	0	1
Parent education	Schooling level of the (biological or non-biological) parent who is living with the child, distinguishing between "never went to school", "not graduate from high school", "high school graduate", "graduated from college or a university", "professional training beyond a four-year college", coded as 1 to 5. We considering only the education of the father if both parents are in the household.	3.58	2.08	1	5
Parents age	Mean value of the age (years) of the parents (biological or non-biological) living with the child.	40.14	13.64	33	75
Parent occupation manager	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. “Doesn't work without being disables” is the reference group	0.11	0.13	0	1
Parent occupation professional/technical	“	0.09	0.22	0	1
Parent occupation office or sales worker	“	0.25	0.29	0	1
Parent occupation manual	“	0.21	0.30	0	1
Parent occupation military or security	“	0.08	0.12	0	1
Parent occupation farm or fishery	“	0.04	0.09	0	1
Parent occupation retired	“	0.06	0.10	0	1
Parent occupation other	“	0.13	0.17	0	1
Protective factors					
Relationship with teachers	Dummy taking value one if the respondent reports to have trouble getting along with teachers at least about once a week, since the beginning of the school year.	0.15	0.35	0	1
Social exclusion	Response to the question: "How much do you feel that adults care about you, coded as 1= very much, 2= quite a bit, 3= somewhat, 4= very little, 5= not at all	2.26	1.80	1	5
School attachment	Composite score of three items derived from the questions: "How much do you agree or disagree that: (a) You feel close to people at your school, (b) you feel like you are part of your school, (c) you are happy to be at your school", all coded as 1= strongly agree, 2= agree, 3=neither agree nor disagree, 4= disagree, 5= strongly disagree. (Crombach-alpha =0.75).	2.57	1.75	1	5
Parental care	Dummy taking value one if the respondent reports that the (biological or non-biological) parent that is living with her/him or at least one of the parents if both are in the household cares very much about her/him	0.65	0.35	0	1
Residential neighborhood variables					
Residential building quality	Interviewer response to the question "How well kept is the building in which the respondent lives", coded as 1= very poorly kept (needs major repairs), 2= poorly kept (needs minor repairs), 3=	2.96	1.85	1	4

	fairly well kept (needs cosmetic work), 4= very well kept.				
Neighborhood safety	Dummy variable taking value if the interviewer felt concerned for his/her safety when he/she went to the respondent's home.	0.52	0.57	0	1
Residential area suburban	Residential area type dummies: interviewer's description of the immediate area or street (one block, both sides) where the respondent lives. Rural area is the reference group.	0.32	0.39	0	1
Residential area urban - residential only	“	0.18	0.21	0	1
residential area industrial properties - mostly wholesale	“”	0.13	0.18	0	1
residential area other type	“	0.19	0.25	0	1
Contextual effects	Average values of all the control variables over the respondent's direct friends (peer-group characteristics).				

Table 1: Correlation between individual, network and peer group-level variables

Variable	(1)	(2)
Parental education	-0.1996	0.0725
	(0.3417)	(0.1198)
Parental care	0.1562	-0.1662
	(0.0631)	(0.2217)
Mathematics score	-0.1819	0.0699
	(0.2042)	(0.0755)
Motivation in education	-0.0896	0.1546
	(0.2577)	(0.1869)
School attachment	0.0725	0.0499
	(0.0993)	(0.0763)
Social exclusion	0.0317	-0.0901
	(0.0341)	(0.1008)
Individual socio-demographic variables	yes	yes
Family background variables	yes	yes
Protective factors	yes	yes
Residential neighborhood variables	yes	yes
Contextual effects	yes	yes
School dummies	yes	yes
Network fixed effects	yes	yes

Notes:
- (1): Estimated correlation coefficients between observables and residuals from network fixed-effects OLS regressions
- (2): Network fixed-effects OLS estimates are reported
- Standard errors (in parentheses) are reported
- Control variables are those listed in Table A.1
- Regressions are weighted to population proportions
- None of the coefficients is statistically significant at any conventional level

Table 2: Maximum likelihood estimation results
Dependent variable: delinquency index

Variable	All crimes	Type 1	Type 2	Type 3
Conformism / peer effects (ϕ)	0.0612**	0.0688**	0.0499**	0.0079**
	(0.0305)	(0.0320)	(0.0241)	(0.0035)
Individual socio-demographic variables	yes	yes	yes	yes
Family background variables	yes	yes	yes	yes
Protective factors	yes	yes	yes	yes
Residential neighborhood variables	yes	yes	yes	yes
Contextual effects	yes	yes	yes	yes
School dummies	yes	yes	yes	yes
Network fixed effects	yes	yes	yes	yes
pseudo-R^2	0.4766	0.4915	0.4111	0.4599

Notes:
- Estimated coefficients and standard errors (in parentheses) are reported
- Estimation using SpaceStat v1.93 (Anselin, 1995).
- Control variables are those listed in Table A.1
- ** indicates statistical significance at the 5 percent level