Centre for Research and Analysis of Migration

CReAM

Discussion Paper Series

CDP No 04/12

Peer Effects on Criminal Behavior. Evidence from the homeless

Lucia Corno

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

CReAM Discussion Paper No 04/12

Peer Effects on Criminal Behavior. Evidence from the homeless

Lucia Corno*

* University College London and Centre for Research and Analysis of Migration (CReAM)

Non-Technical Abstract

This paper investigates the influence of peers on criminal behavior, using original data I collected by interviewing homeless people in Milan. Information on friends' names was elicited, which allows to map each respondent's network. Each individual was also asked to report his criminal status prior to becoming homeless. To estimate the causal effects of network size and of the share of criminal friends on (subsequent) criminal behavior, I rely on two instruments. The first is the share of rainy days since the individual has become homeless: rainfall fosters concentration of homeless individuals in sheltered places and increases the probability of meetings. The second instrument is the fraction of inmates released by Milan's authorities during one's period as homeless, which affects the supply of criminal potential friends. I find that the probability of arrest decreases by 16 percentage points with the network size, but it increases by 20 percentage points with the share of criminal friends in the group.

Keywords: Peer effects, crime, homeless.

JEL Classification: J0; K42.

Centre for Research and Analysis of Migration
Department of Economics, Drayton House, 30 Gordon Street, London WC1H 0AX
Telephone Number: +44 (0)20 7679 5888
Facsimile Number: +44 (0)20 7679 1068

Peer Effects on Criminal Behavior. Evidence from the homeless

Lucia Corno*

February 5, 2012

Abstract

This paper investigates the influence of peers on criminal behavior, using original data I collected by interviewing homeless people in Milan. Information on friends' names was elicited, which allows to map each respondent's network. Each individual was also asked to report his criminal status prior to becoming homeless. To estimate the causal effects of network size and of the share of criminal friends on (subsequent) criminal behavior, I rely on two instruments. The first is the share of rainy days since the individual has become homeless: rainfall fosters concentration of homeless individuals in sheltered places and increases the probability of meetings. The second instrument is the fraction of inmates released by Milan's authorities during one's period as homeless, which affects the supply of criminal potential friends. I find that the probability of arrest decreases by 16 percentage points with the network size, but it increases by 20 percentage points with the share of criminal friends in the group.

JEL Classification: J0; K42
Keywords: Peer effects, crime, homeless

*University College London and Centre for Research and Analysis of Migration (CReAM), Department of Economics, 30 Gordon Street, London WC1H OA. Phone: +44-(0)2-076795451. Fax: +44-(0)2-076791068. Email: l.corno@ucl.ac.uk. This project required the collaboration of many people. First of all, I am deeply indebted with Michela Braga, my co-worker during all stages of the project. The data collection benefits from the help of many volunteers, in particular from the Italian Red Cross, Caritas and Opera Cardinal Ferrari. I wish to thank Eliana La Ferrara and Edward Miguel for their valuable comments and suggestions. I also benefit from the inputs by Emilio Calvano, Stefano Dellavigna, Gunther Fink, Masayuri Kudamatsu, Enrico Moretti, Tommaso Nannicini, Steve Pischke, Imran Rasul, David Stromberg and seminar participants at Berkeley, Bocconi University, University College London and IIES at Stockholm University. I thank the Regional Agency for the Environmental Protection and the Meteorological Service of the Military Aereonautics for making rainfall data available to me and correctional facilities of S.Vittore, Bollate and Opera to provide me administrative data on inmates. I thank Marzio Barbaglio for sharing data on police force deployment in Italy. Financial support from the Empirical Research in Economics (ERE) and Fondazione Rodolfo De Benedetti (FRdB) is gratefully acknowledged. All errors are my own.

"A closed door may make feel you safe, but an open door may save your life. Let's open the doors!"

Ambrogio, Homeless in Milan

1 Introduction

Homelessness is a critical problem among the urban poor. Their extreme social exclusion and isolation reinforce the role of peers in shaping individual behavior. Among homeless people, peers are the main source of information about potential jobs, shelter locations and welfare programs and they might also provide informal insurance against idiosyncratic shocks. Randomized experiments in developing countries also show that peers have a considerable influence on the decisions of the poorest.[1]

Peer effects have been widely documented in many contexts, from educational choices to labor market outcomes. The role of social interactions is particularly important in the criminal sector, where informal networks compensate for the lack of formal institutions in gaining knowledge and criminal skills.[2] However, the empirical evidence documenting causal effects of peer behavior on individual criminal decisions is not convincing. The main difficulties are related to endogenous group formation - to the extent that unobservable characteristics affecting the network's composition are likely to be correlated with unobservables influencing the decision to engage in crime - and reflection - since in a peer group everyone's behavior affects the others' behavior and it is problematic to disentangle the individual's behavior from that of the reference group (Manski, 1993). Finally, the scarcity of individual data publicly available on social interactions limits the identification of network boundaries only at a quite aggregate level, potentially leading to attenuation bias from measurement errors (Kling, Ludwig, Katz, 2005).

The goal of this paper is to estimate the causal impact of peer characteristics and of network size on criminal behavior among the homeless. Specifically, it addresses the following questions: how does individual criminal behavior change with peer characteristics? Is criminal behavior influenced by the size of one's social network? I test these questions using an original survey I conducted by interviewing homeless people in Italy. Investigating the reasons for criminal acts is particularly

[1] For example, Bobonis and Finan (2009) study neighborhood effects on children's school enrollment using experimental evidence from the Mexican PROGRESA program. They found that peers have a considerable influence on the enrollment decisions of children from poorer households. Godlonton and Thornton (2010) suggest that peer effects in learning HIV results might be stronger in rural settings where other forms of communication (such as television, radio or newspaper) are more limited.

[2] The concept that criminal behavior is a *learned* behavior has been emphasized by Sutherland (1947) and, subsequently highlighted, among others, by Bayer, Hjalmarsson and Pozen (2009); Calvo-Armengol and Zenou (2004); Case and Katz (2001); Cook and Gross (1996); Glaeser, Sacerdote and Sheinkman (1996); Kling and Ludwig (2007).

interesting among the homeless. In the last few years, the face of homelessness has changed substantially from the classical stereotype. Homeless people represent, on average, a group of urban poor who simultaneously experienced economic or family related shocks, or former middle class and business owners who can no longer meet their financial obligations. This creates a population forced to live in poverty-stricken conditions, where criminal behaviors became over-represented because they perceived benefits outweigh the cost of potential punishment (Becker, 1968).

The novelty of the paper is two-fold. First, the data come from the first representative survey in Europe among the homeless, collected by the author in January 2008 in Milan, Italy. The survey is particularly suitable to study peer effects because it includes names and surnames of each respondent and those of their five best homeless friends. Thus, I am able to construct the geometric structure of the friendship networks and to precisely identify a close set of individual's peer group. Information is also available on each respondent's criminal status prior to becoming homeless. The analysis is based on 561 sheltered and unsheltered homeless individuals interviewed in one single night, with a high proportion of former inmates in the sample (29.6%). Second, the paper estimates the causal effects of network size and of the share of criminal friends on (subsequent) criminal behavior in a way that deals with the non-random matching of individuals to their peers. Specifically, the identification strategy exploits two instruments. The first one is the share of rainy days during one's period as homeless to instrument the network size: rainfall fosters concentration of homeless people in sheltered places (i.e. bridges, underground, train stations) and increases the probability of meetings. The second instrument is the share of inmates released by Milan's authorities during one's homeless window and it is used to correct for the self-selection of past criminals and potential criminals in the same network.

To motivate the empirical analysis, the paper offers a simple theoretical framework to model the choice between committing or not committing a crime, depending on the fraction of criminal friends and on the total number of friends in the network. The model shows that it exists a unique equilibrium where the individual probability to commit a crime is equal to the fraction of criminal friends in the group. The theoretical framework provides testable predictions for the empirical section.

I start the empirical analysis by investigating the effect of the network size, measured by the number of best friends, and the role of peers, captured by the share of friends who have already been in prison *before* homelessness, on the individual probability of imprisonment *during* homelessness. I consider the different timing of imprisonment between the respondent and his friends (*during* and *before* being homeless, respectively) to disentangle the direction of causality from peers to individual's behavior. Following Bayer, Hjalmarsson and Pozen (2009), I assume that peer effects operate through the influence of peer characteristics. To address the endogeneity in the number of friends, I instrument the size of the network with the fraction of rainy days in one's period as homeless. Rainfall shocks randomly allocate individuals on the street, causing concentration in sheltered places (i.e. bridges, underground, train stations) and a consequent higher probability of new friendships. Indeed, changes in the fraction of rainy days strongly and positively predict the number of friends. Ceteris paribus, a one standard deviation increase in the fraction of rainy days during one's period as homeless is associated with having 0.33 more friends. The key identification assumption here is that the fraction of rainy days in one's period as homeless does not directly influence the probability of imprisonment. Detailed checks for the validity of the exclusion restriction are provided in the paper. To correct for the self-selection of past criminals and potential criminals in the same network, I instrument the share of criminal peers in an individual's social network with the fraction of inmates released by Milan's authorities during one's period as homeless. In this case, the underlying assumption is that exogenous policies driving inmates' outflow increase the supply of criminal potential friends, without directly affecting individual criminal behavior. Ceteris paribus, a one standard deviation increase in the number of inmates released increases the share of criminal peers by 3.9 percentage points. Different dates of arrival on the street provide the necessary individual variation in the instruments used, which is plausibly as good as random.

Based on the instrumental variable estimates, the likelihood of imprisonment during a homeless window *decreases* by 16 percentage points on average with an additional friend. The interpretation of this finding is consistent with the notion that friends represent a source of mutual insurance. Homeless individuals with more friends have greater chances to survive on the street without committing crime because their idiosyncratic shocks are shared among a higher number of individuals. Results also indicate the presence of significant peer effects in the probability to commit criminal acts: a one standard deviation increase in the share of peers who served previous jail sentence *increases* the probability of incarceration for an individual with no prior criminal experience by 20 percentage points.

The reminder of the paper is organized as follows. Section 2 reviews the literature on social networks and crime and provides a background on the data collected among homeless people. In section 3, a simple theoretical framework is offered to highlight the relationship between individual criminal behavior, criminal friends and network size. Section 4 describes the research design for the data collection and section 5 illustrates data and descriptive statistics. In section 6, I present the identification strategy. Sections 7-8 show the main results and robustness checks, section 9 concludes and presents some policy recommendations.

2 Background

2.1 Previous works on social networks

This paper fits into the literature investigating the role of social interactions in shaping individual criminal behavior.[3] Starting from Sutherland (1947), criminological and economic studies recognized the potential channels through which peers may influence criminal activity: by sharing information (Calvò-Armengol and Zenou, 2004), by recruiting young criminals (Reiss, 1988), by transferring skills or by creating a sense of invulnerability (Glaeser et al. 1996), by affecting the social stigma associated with the illegal act (Cook and Gross, 1996). However, few papers have been convincing in measuring the causal impact of peers on individual criminal decision.

There are major difficulties in estimating the influence of social interactions on criminal behavior. First, the scarcity of individual data publicly available limits the identification of social network boundaries at the aggregate level. Several contributions study neighborhood effects on criminal acts. Case and Katz (1991) find that residence in a neighborhood where 10% more of the youth are involved in crime raises the individual's probability to become a delinquent by 2.3%. Subsequent research examines the role of neighborhoods in predicting criminal activities, by relying on particular randomized events (Katz, Kling, Leibman, 2001; Ludwig, Duncan and Hirschfield 2001; Kling, Ludwig and Katz, 2005; Kling and Ludwig, 2007). Glaeser et al. (1996) provide an index to measure the importance of social interactions in crime across cities in US and they found that the amount of social interactions changes depending on the severity of crime. More recent literature considers correctional facility boundaries to study peer effects on post released behavior. Bayer et al. (2009) analyze the influence that juvenile offenders serving time in the same correctional facility has on each other's subsequent criminal behavior. However, being part of the same neighborhood or of the same correctional facility does not necessarily imply social interactions among these individuals. On this respect, the data collected for the present paper are unique, since I am able to delineate a close set

[3]The influence of peers has been widely documented in many other contexts. Among others, peer effects on educational outcomes have been studied by Sacerdote (2001), Calvò-Armengol, Patacchini and Zenou (2009); on educational choices by De Giorgi, Pellizzari and Redaelli (2010); on productivity at the workplace by Mas and Moretti (2009), Bandiera, Barankay and Rasul (2010); on labor market outcomes by Munshi (2003), Beaman (2008); on welfare participation by Bertrand, Luttmer, Mullainathan (2000); on technology adoption by Bandiera, Rasul (2006), Conley, Udry (2010).

of individual's peer group, identified by actual friends' nomination.[4]

The other difficulties in identifying the effect of social interactions are related to endogeneity - to the extent that unobservable characteristics affecting the likelihood to have a friend are likely to be correlated with unobservables influencing the decision to engage in crime (i.e. peers self-selection and common group effects) - and reflection - since in a peer group everyone's behavior affects the others' behavior and it is problematic to disentangle individual's behavior from the one of the reference group (Manski, 1993). To address this issue, recent research has relied on a specific randomized social experiment, the Moving To Opportunity in Boston, which randomly allocates housing vouchers from high to low poverty neighborhoods to examine the impact of newly allocated neighborhoods on criminal activities. Although the randomization present in these experiments is ideal, relying exclusively on such events severely limits the setting where peer effects can be studied (Bayer et al., 2009). By including facility and facility-by-prior-offense fixed effects to control for the non-random assignment to facilities, Bayer et al. (2009) find that peers, with a history of committing a particular crime, increase the probability that an individual who has already committed the same type of crime recidivates with that crime. In contrast, there is no evidence of such peer effects for individuals with no prior experience in a given crime category. In this paper, I propose an instrumental variable identification strategy, by relying on individual exogenous variation provided by different dates of arrival on the street/shelter.

Furthermore, prior empirical research mainly tests the role of criminal peers on individual criminal behavior, while the effect of the total network size has been overlooked. Using the total number of nominated friends as a proxy for the size of the social network, this paper investigates whether individuals in larger or smaller networks are more or less likely to commit illegal offenses.

2.2 Review of the existing data collection

Homelessness is a public policy issue in many countries. However, the lack of reliable data has limited economic research and effective strategies to prevent and reduce the phenomenon. So far, only a few countries include official statistics on homelessness in the general population census or have developed *ad hoc* methodologies and procedures to regularly count homeless people. In the US, the institution that regularly carries out counts of homeless individuals is the Department of Housing and Urban Development (HUD). Since 1984, the HUD requires counts of this population every two years on a

[4] In the economic literature, only a few surveys recorded information on individual's or household's social networks, based on friends' nomination: the National Longitudinal Study of Adolescent Health (Patacchini and Zenou, 2012); a panel household survey conducted in Tanzania (Dercon and De Weerdt, 2006; De Weerdt and Fafchamps, 2011); a survey data from rural Philippines (Fafchamps and Lund, 2003; Fafchamps and Gubert, 2006).

national sample of 80 communities in different geographical areas. The HUD's most recent estimates indicate 754,000 persons in the US living in emergency shelters, transitional housing, and on the street on any given night. For the first time in 1990, the US Census Bureau included in the Decennial Census the collection of data on homeless people into the general population census in five US Cities (Chicago, Los Angeles, New Orleans, New York and Phoenix). Up to now, official countrywide data are not available for Europe, but, in a few countries, counts of homeless population have been conducted by local agencies or NGOs. Besides the Milan Homeless Survey (MHS) conducted by the author in January 2008, in Italy only two other attempts have been made to carry out a data collection on homeless people. The first survey was jointly led by the Commission on Social Exclusion at the Ministry of Social Affairs, and the Padua Zancan Foundation. This survey was aimed at delineating the characteristics of people without a dwelling and to estimate their number throughout the entire national territory. The final figure accounts for a population of 17,000 homeless people in Italy, with a higher concentration in the bigger municipalities. At a regional level, in Veneto, the University of Padua and the Regional Observatory for the Protection and Promotion of the Person conducted a survey in seven Italian cities in December 2004, on a sample of about 140 homeless people in shelters and on the street, with the aim of gathering their socio-demographic characteristics.

Research on homelessness mainly analyzes the causes of the substantial increase in the incidence of homelessness during 1980s in the US, using data provided by the HUD (Quigley et al., 2001; Honig and Filer, 1993; O' Flaherty, 1996; Quigley, Raphael and Smolenski, 2001) or by the US Census Bureau (Burt, 1990). A common result among these studies is that variation in homelessness arises from changed circumstances in the housing market and in the income distribution. Specifically, tougher housing markets are positively associated with higher levels of homelessness. While previous works used intercity aggregate data, this is the first paper using micro-level data to study homelessness from an economic perspective. Another strand of the literature investigates homelessness duration and it underlines that homeless population experiences temporary but recurrent spells of homelessness (Piliavin, Wright, Mare and Westerfelt, 1994). Homeless spells are generally longer for those with history both of drugs and alcohol abuse, while having received government benefits in the past seems to have a controversial effect on the average length of a homeless spell (Allgood and Warren, 2003; Braga and Corno, forthcoming).

3 A simple theoretical framework

The goal of this section is to provide testable predictions on the impact of peers affecting the individual decision to commit a crime. I investigate how homeless individuals react to an exogenous change in the number of friends and in the fraction of criminal friends.

Consider a continuum of individuals. Individuals are randomly connected in a network of social interactions. Suppose that in each network, a fraction α is criminal and the reminder fraction $1-\alpha$ is non-criminal. This fraction is assumed to be equal to the fraction of criminals in the general population. We assume that individuals randomly become friends to any other individuals in the network, who can be criminal or not. N_i represents the total number of friends for the individual i and the number of criminal friends, $N_{i,CR}$, is given by:

$$N_{i,CR} = \alpha N_i$$

The total number of non-criminal friends, $N_{i,NCR}$, is then defined as:

$$N_{i,NCR} = (1-\alpha)N_i$$

Each agent bears a cost of committing a crime, C. This cost is negatively correlated with α, the fraction of criminal friends: (i) in a network of criminals, each person can exert less criminal effort for the same criminal payoff compared to operating alone; (ii) criminals share their knowledge about delinquent behavior between their friends and since there are no formal ways of learning how to become a criminal, the most efficient way is through criminal friends. To become a criminal, a person must also be inclined toward illegal activity (Patacchini and Zenou, 2012). Thus, suppose that the Nature draws for each agent a propensity to commit a crime $\theta_i \in R_+$ from a CDF F with continuos density. I assume that the agent's realization of θ_i is his private information. The cost of committing a crime C, is negatively correlated with the individual criminal propensity, θ_i. A higher θ_i means a lower psychological cost to commit illegal acts.

Following the seminal work of Becker (1968), I assume that individuals trade off the costs and benefits of delinquent activities to take their decision between committing or not committing a crime. A person commits an offense if the expected utility to him exceeds the utility he could get by using his time in other activities. The expected utility from committing crime is:

$$EU_{i,CR} = pU(Arrest) + (1-p)U(NonArrest) - C(\alpha, \theta_i)$$

where p is the probability of conviction and it depends only on the level of police efficiency. As explained above, the cost of committing crimes, C, is negatively correlated with the fraction of criminal friends, $\alpha : C_\alpha = \frac{\partial C}{\partial \alpha} \leqslant 0$ and with the criminal propensity, θ_i, so that $C_{\theta_i} = \frac{\partial C}{\partial \theta_i} \leqslant 0$. Furthermore, assume that $C_{\alpha\alpha} = \frac{\partial C}{\partial \alpha} > 0$ and $C_{\alpha\theta_i} = \frac{\partial C}{\partial \alpha \partial \theta_i} < 0$.[5] The expected utility from non committing a crime is equal to the outside option:

$$U_{i,NCR} = W_o(N_i);$$

where the total number of friends, N_i positively affects $U_{i,NCR}$. It follows that $\frac{\partial U_{i,NCR}}{\partial N_i} > 0$.

The key idea behind this assumption is the plausible presence of a mutual insurance among friends. Friends help to smooth income shocks by transferring resources after the shock is realized. For example, it is plausible that not all the homeless in the same group have a job at the same time. Having idiosyncratic risk creates solidarity mechanisms. Thus, risk-sharing benefits depend positively on the number of total friends N_i. [6]

Definition 1: An individual commit a crime if and only if

$$EU_{i,CR} \geq U_{i,NCR}$$

that is when:

$$pU(Arrest) + (1-p)U(NonArrest) - C(\alpha, \theta_i) \geq W_o(N_i).$$

An individual with a realization θ_i will decide to commit crime if and only if $\theta_i > \overset{\sim}{\theta}$ where $\overset{\sim}{\theta}(\alpha; N_i)$ is defined as:

$$pU(Arrest) + (1-p)U(NonArrest) - C(\alpha, \overset{\sim}{\theta}) = W_o(N_i). \tag{1}$$

From equation (1), we can evaluate a change in the decision to be involved in criminal activities in response to an increase in the total number of friends, N_i and to an increase in the fraction of

[5] Such conditions are made for convenience since they guarantee a unique equilibrium, without affecting the relative comparative static.

[6] Similarly, Bramoullé and Kranton (2007) build a theoretical model of risk sharing between pairs of agents where the risk sharing benefits depend on the number of agents in each village.

criminal friends α. Standard applications of the *implicit function theorem* lead to:

$$\widetilde{\theta}_{\alpha} = \frac{\partial \widetilde{\theta}(\alpha, N_i)}{\partial \alpha} = -\frac{-\,C_{\alpha}}{-\,C_{\widetilde{\theta}}} = -\frac{C_{\alpha}}{C_{\widetilde{\theta}}} < 0$$

and

$$\widetilde{\theta}_{N_i} = \frac{\partial \widetilde{\theta}(\alpha, N_i)}{\partial N_i} = -\frac{-W_{o,N_i}}{-\,C_{\widetilde{\theta}}} = -\frac{W_{o,N_i}}{-\,C_{\widetilde{\theta}}} = \frac{W_{o,N_i}}{C_{\widetilde{\theta}}} > 0.$$

These results drive to the following predictions:

Prediction 1: An increase in the fraction of criminal friends α will decrease $\widetilde{\theta}$. A lower threshold, $\widetilde{\theta}$, has the effect of reducing the cost of crime, and consequently to increase the expected utility of committing a crime. This implies that the probability to commit a crime increases with a higher share of criminal friends in the network.

Prediction 2: An increase in the total number of friends, N_i will increase $\widetilde{\theta}$. In this case, a higher threshold, $\widetilde{\theta}$, has the effect of increasing the cost of crime, and consequently to decrease the expected utility of committing a crime. This implies that the probability to commit a crime decreases with a higher number of friends.

The proportion of individuals with $\theta_i > \widetilde{\theta}$, who will then choose to commit a crime, is given by $1 - F(\widetilde{\theta}(\alpha; N_i))$[7]. The equilibrium of the model is then:

$$\Pr(\theta_i > \widetilde{\theta}) = 1 - F(\widetilde{\theta}(\alpha; N_i)) = \alpha.$$

The existence of an equilibrium follows directly from the Brouwer's Fixed Point Theorem. To see that this point is necessarily unique, let $g(\alpha) = 1 - F(\widetilde{\theta}(\alpha; N_i))$ and consider the sign of the first and the second derivatives. Differentiating $g(\alpha)$, we get: $g_{\alpha} = -F_{\widetilde{\theta}}\widetilde{\theta}_{\alpha} > 0$ and $g_{\alpha\alpha} = -F_{\widetilde{\theta}\widetilde{\theta}}\widetilde{\theta}_{\alpha} - F_{\widetilde{\theta}}\widetilde{\theta}_{\alpha\alpha} > 0$.[8] This implies that $g(\alpha) = 1 - F(\widetilde{\theta}(\alpha; N_i))$ crosses $g(\alpha) = \alpha$ only in one point and that point is unique.

[7] This result is closely related to a recent work by Anwar and Fang (2006), developed in the literature on racial bias in motor vehicle searches.

[8] The second derivative of $\widetilde{\theta}$ with respect to α is equal to: $\widetilde{\theta}_{\alpha\alpha} = \frac{-C_{\alpha\alpha}C_{\widetilde{\theta}} - C_{\alpha\widetilde{\theta}}C_{\alpha}}{(C_{\widetilde{\theta}})^2} < 0$

4 Survey design and data

4.1 A survey among the homeless

The data used in this paper come from an innovative and representative survey among homeless people, managed by the author in January 2008 in Milan, Italy. The project involved about 350 volunteers, recruited among service providers to the homeless, but also among students and private citizens, thanks to the substantial interest received by the project from local media and newspapers.

A survey among homeless people involves many challenges. First, it is difficult to clearly define the target population. Our reference population includes all persons who reside in (i) places not meant for human habitation, such as cars, parks, sidewalks, abandoned buildings (unsheltered homeless); (ii) emergency shelters (sheltered homeless); (iii) people living in disused areas/shacks/slums.[9] The second challenge arises since it is very difficult to provide reliable estimates on the number of homeless individuals in a city and to conduct an accurate survey. The Milan Homeless Survey 2008 (MHS 2008) includes two major phases: 1) a point-in-time count; and 2) a comprehensive survey via trained interviews.

The *Point in time survey* or the *S - Night approach (Shelter and Street Night)* aims at identifying the number of homeless sleeping on the street, in shelters and in slums contemporaneously in one reference night in the whole town. As such, it ensures a minimal double counting. Homeless people are indeed both territorially mobile and likely to enter into and exit out of the homeless state, and the risk to count and to interview the same person twice is therefore very high.[10] Our reference night for the count was January 14^{th}, 2008. The main drawback of the point-in-time method is the risk of missing homeless hidden from public view during late-night hours. We applied some efforts to overcome this criticism and to have the most reliable estimate. We divided Milan into 65 smaller areas, following the main roads, so that a team of 3-4 enumerators could reasonably cover them during the night of the count. Surveyors were asked to walk every street and other public places in their assigned area. To reduce the risk of skipping some streets, we provided the enumerators

[9] The third category represents a very peculiar feature of the Italian context and refers to illegal settlements which are mostly inhabited by (irregular) immigrants and gypsies. The choice to include them in our definition of homelessness was motivated by the fact that these arrangements can be classified as inadequate or insecure housing situations.

[10] Among the most recent approach proposed to count the homeless population, is the *capture-recapture* method. This method calculates the total homeless population from the sum of the population actually observed and an estimate of the unobserved population, by calculating the number of people not caught in either sample. A limitation of this method consists in estimating the homeless population during an entire year. Therefore, it assumes that all individuals identified as homeless remain homeless for the full year (Fisher et al. 1994). Brent (2007) and Braga and Corno (forthcoming) provide a detailed overview of different methods used to count homeless populations.

with enlarged maps of their target area, and we defined in advance the itinerary to be followed, by writing down the complete list of all streets in each area. We established some criteria for the count: closed tents and closed paperboard dwellings have been counted as for one homeless person, while for abandoned cars/caravans enumerators tried to understand how many persons were sleeping there. To be sure all enumerators started the count at the same time, they met one hour before the kick-off in five strategic places. There, they also collected useful materials for the night (i.e. torches, food, beverages and notebooks). Besides counting, volunteers have two additional tasks: (i) to report a homeless person's location as precisely as possible, by describing the road, the closest civic number but also the sleeping place (i.e. Sarfatti road close to number 25 on a bench in front of Bocconi University); (ii) to detect some observable characteristics, such as the ethnic group, gender and estimated age: this was a key information to test for a potential sample selection. Volunteers paired these statistical activities with hot beverages and food distribution.

In the meantime, a team of volunteers collected information on the number, names, gender, age and nationality of the homeless living in the emergency shelters in the city.

The procedure for counting people living in slums was not straightforward. Slums in Milan are settlements made by prefabricated materials or set up in disused barracks, where people (mainly gypsies) are generally monitored by the municipal police. During the three months prior to the survey, project leaders visited the slums in the target group, to identify the typology of the village (authorized/unauthorized), the type of ethnic group and the number of people living in each area. During the field visits, we asked for the permission to interview people in the slums and we announced the date of the survey. On the night of the count, enumerators only checked dimensions and locations of pre-identified slums. The average duration of the count was about 3 hours, from 10 pm. to 1 am.

The count was necessary to have a precise idea on the phenomenon's dimension and to construct a census from which we randomly selected a sample of respondents. Questionnaires to the unsheltered homeless were performed on the following night, January 15^{th}, while we surveyed people who were sleeping in shelters and in slums on January 16^{th} and 19^{th}, respectively.[11] [12] The whole data collection was then completed in a single week to minimize sample attrition. The survey involved a total of 75 interviewers out of 350 volunteers. To minimize answer bias, we intensively trained the interviewers and we recruited the same interviewers for all the three nights. On the street, we tried to have the full census of the homeless counted by sending back enumerators to the locations identified during the

[11] Interviews in slums have been done on Saturday afternoon. We decided not to go in slums during the night for security reasons.

[12] The questionnaire is available upon request.

count. Sheltered homeless were randomly sampled from the population on the basis of the shelter's dimension. We created a random sample proportional to the shelter dimension by over-sampling small shelters and under-sampling big shelters. We agreed in advance on the best time to run interviews with each shelter's manager. Among 25 shelters, four refused to participate and one had no guests at the time of the survey. Some interviews were conducted directly by shelter managers. Finally, slums were sampled through a stratified random sample method, based on geographic location, typology and dimension. More specifically, we stratified them accordingly to city administrative division (9 areas), official area classification (authorized/unauthorized, shacks, abandoned buildings) and area's dimension.[13] We selected a total of 12 out of 56 slums. Within each selected area, we randomly extracted respondents. During the interviews, volunteers also distributed napkins and kids' clothes to the households in the slums. To preserve enumerators' safety, we informed the police about the initiative without directly involve it.

A potential drawback in doing the count and the interviews on two different days (even if in very close proximity) is the attrition rate, since people counted could have moved the day after. To control for the fact that the homeless counted were the same as those interviewed, we included as the first question in the survey "Did you sleep here last night?" and if no "Where did you sleep?". We cross checked this information with the homeless locations recorded during the count.

Self-reported answers can be biased for many reasons. This might be particularly true with surveys among homeless people: they can be drunk during the survey and they are more likely to be mentally ill compared to the general population. To address this drawback, we asked to the enumerators from the Red Cross to fill up a one page questionnaire regarding respondent's condition at the end of the survey. We did not consider questionnaires conducted with drunk or mentally-ill homeless persons (2%). As incentive, enumerators distributed grocery vouchers to the respondents who fully completed the questionnaire. The questionnaire was written in Italian and translated into Romanian and English. The average time for an interview was about 30 minutes.

The count and the interviews were not conducted on the same day for two main reasons. First, it is not feasible to interview people during a one-night count. During the count, enumerators minutely checked the presence of homeless by walking along all streets in Milan and there would not be remaining time to also select and interview them. Second, while it is optimal to conduct a late-night count (from around midnight to 3 a.m.) to maximize the probability to observe more

[13] Slum's dimension was classified as follows: "small" if inhabitants numbered less than 30, "medium" if inhabitants numbered between 30 and 100, "big" if inhabitants numbered more than 100.

visible people sleeping outside, the ideal time for interviews is around 9 p.m. when they are already settled down, but still awake and able to talk. The survey took place in January when the average daily temperature is the lowest in Milan and shelters are likely to be at peak capacity: it is easier to count people in shelters than on the street and conducting the count on a night when shelters are most full will likely lead to the most accurate count. Counting and interviewing people sleeping in open locations during the winter months may also lead to a more realistic picture of chronically unsheltered homeless. Furthermore, to facilitate the identification of homeless people and to reduce the likelihood of the surveyors being overwhelmed by potential respondents, we chose a day of the week with less pedestrian traffic (Monday night).

4.2 Additional Data Sources

The paper exploits two additional data sources. First, I use rainfall data to proxy the size of the friendship network. Rainfall data come from the Regional Agency for the Environmental Protection (ARPA) and from the Meteorological Department of the Military Aeronautics. Daily rainfall data has been collected from 1960 to 2008. I use information on rainfall from six weather stations within Milan municipality, but some stations lack data for some days per month.[14] I calculate the average number of rainy days among weather stations as a proxy for daily rainfall in Milan.[15]

Second, I assembled administrative data on inmates from the Statistical Offices of the three correctional facilities in Milan.[16] Specifically, these data include the number of criminals arrested and released, and the fraction of inmates without a residence at the moment of the arrest, by month from 1993 to 2008.[17] To proxy the number of inmates released in Milan, I compute the monthly sum of inmates released among correctional facilities.

5 Descriptive statistics

The population of homeless in Milan accounts for 3860 homeless: 408 unsheltered homeless, 1152 sheltered homeless and about 2300 adults (older than 16 years old) in slums. This approximately represents 0.3% of the total population in Milan.

[Insert table 1]

[14] The six weather stations in Milan are located in Lambrate, Parco Nord, Zavattari, Confalonieri, Juvara, Brera.
[15] As defined by the ARPA, a rainy day records at least one millimeter of rainfall during 24 hours.
[16] There are three correctional facilities in Milan: S. Vittore, Opera, Bollate.
[17] Data on inmates arrested and released in Bollate are available only from 2000, when this correctional facility was open.

Table 1 shows the percentage of those who did not participate in the survey, by place of interview. Among the homeless counted on the street on January 14^{th} 2008, we interviewed about 34.6%, 12% refused to answer, 16.4% were already sleeping at the time of the interview and 21% were not found. Due to time constraints, we did not send enumerators in 16% of the identified locations.[18] In shelters, we randomly sampled 500 individuals out of 1152 and we interviewed 420 of them. While 6.7% of the sample was not in the shelters on the day of the interview, 7.3% was not interviewed for lack of time and about 2% refused to answer. In slums, we randomly selected a sample of 525 individuals out of 2300 and we surveyed 66.5% of the sample, but we did not conduct the questionnaire for about 33.5% of the respondents for lack of time. None among the slum dwellers refused to participate. We dropped a small fraction of bad quality questionnaires, in which the enumerators reported respondents with mental disorders or alcohol-related problems (0.02%). We gathered a final sample of 910 observations. This paper focuses on street and sheltered homeless, ending up with 561 observations.[19] To have a better insight on the magnitude of a possible sample selection, I compare data on gender and age of the homeless counted with those of the homeless actually interviewed. The percentage of women interviewed is exactly equal to the percentage of women in the total homeless population (10%) and the percentage of homeless older than 35 years is very similar considering the sample or the total population (72.1% versus 72.4 %). This comparison provides some confidence about the random nature of the sample we are analyzing.[20]

The most common nationalities in the sample are Italians (44.3%), Moroccans (16.52%), Romanians (11.6%) and Egyptians (3.57%).

[Insert figure 1]

Figure 1 reports the spatial distribution of street homeless and shelters' location in Milan. We found a high concentration of street homeless in the centre of the city, in the proximity of train stations (Cadorna Station and Central Station) and at Linate's airport, where every night usually about 15-20 people are sleeping. However, from the inspection of the spatial distribution it emerges that street homeless are almost equally spread within the city. As shown in figure 1, in Milan there are 25 shelters, mainly located in the suburbs.

[18] For example, we decided not to send enumerators in locations recorded as "places with paperboards or abandoned cars, but without individuals".

[19] People who live in slums have a social network structure not comparable with the one of the homeless who live on the street and in shelters. Their social network is mainly represented by the family members they live with. Therefore, I did not include them in the analysis.

[20] Unfortunately, there is no formal way to test a potential sample selection of non-criminal homeless. Criminal homeless might be less likely to participate in the survey. An argument against this hypothesis it that enumerators were Red Cross volunteers in uniform, a clear sign that the survey was totally unrelated to the police.

A first relevant question to answer is: what are the main reasons driving people living on the street? Some of them, such as unemployment and poverty, can be predictable, but others are less intuitive. Results shows that about 33% of immigrants and 24% of Italians cite unemployment as the main cause of their homelessness, either because they lost a job or because they cannot find a job. A breakdown in family relationships, such as a divorce or a death in the family is the main reason for about 36% of Italians, suggesting that family is an important source of insurance against economic and psychological shocks. For foreigners, the second most widespread reason is immigration: at the beginning of their stay in the host country, immigrants have problems related to their limited language proficiency, their scarce knowledge of the Italian welfare system, the labor and the housing markets. Among the other reasons, 9.7% report drug or alcohol abuse and 7% cite previous conviction (not displayed).[21]

[Insert figure 2]

A key feature of the data is that each respondent reports different date of arrival on the street/shelter (day/month/year) and, consequently, different length of their homeless spell.[22] Figure 2 reports the number of homeless people interviewed by the date of arrival (month/year) on the street or in shelter. The average duration of a homeless spell is about 5 years, 7 years for individuals interviewed on the street and about 4 years for those in emergency shelters. The respondent with the longest homeless window has arrived on the street for the first time in 1960.

[Insert table 2]

A section of the questionnaire also investigates their sources of income. Table 2 contains information on the fraction of homeless who declares to have an income, by type of source. The first column shows the figures considering the whole sample of street and sheltered homeless. About 13% of respondents do not have any income. Government subsidies (welfare check, unemployment benefits, disability insurance, pension) together account for about 13.6% and, as expected, this fraction is higher for the Italians.[23] The fraction of homeless who declares earned income (from permanent and occasional jobs) accounts for about 30% of the sample, suggesting a surprisingly high percentage of

[21] For a detailed discussion about the characteristics of the homeless population in Milan see Braga and Corno (forthcoming).

[22] The length of a homeless spell is computed by taking the difference in months between the date of the first time an individual slept on the street/shelter and the date of the survey. In an ideal setting, I would use the date of exit from the homeless status to calculate the end of the homeless spell. Unfortunately, it was impossible to recruit for the survey former homeless people.

[23] To be eligible for welfare checks, individuals must be Italian and resident in Italy, while regular immigrants can benefit from pension, disability and unemployment insurance if they meet the eligibility requirements.

homeless in the labor force. In line with the statistics reporting a breakdown in family relationships as the main cause of homelessness, only about 11% of respondents receive money transfers from family and friends, 14.4% among the immigrants and about 6.8% among the Italians. Table 2 reports that only a few people declare to gain from illegal activities.

[Insert table 3]

However, by investigating criminal behavior from a different perspective, the Milan Homeless Survey 2008 highlights a strong relationship between homelessness and crime. Specifically, the survey investigates whether the respondent has ever been in prison and if this happens before or/and after he slept on the street/shelter for the first time.[24] As reported in table 3, 29.6% of the homeless have been in prison at least once (38.2% of Italians and 21.5% of immigrants). Roughly, 9% declare a period in prison *before* being homeless, 1.8% went to prison *before and during* their homelessness and almost 19% have been arrested only *during* their period as homeless, showing how in extreme poor conditions, crime could become more frequent. Following Locher and Moretti (2004), I use the likelihood of imprisonment as a proxy for criminal behavior throughout the analysis, assuming that a person's probability of conviction is an increasing function of the number of offenses he commits. A limitation of the survey is that it does not elicit information on when the crime was committed and on the type of criminal offenses. To have some insights on the type of crime generally committed by homeless people, I assembled administrative data from correctional facilities in Milan on the criminal acts committed by the inmates who declared "missing residence" when arrested. Typically, these are burglary, robbery, felony and misdemeanor drug offenses, followed by realization of false identification documents and offenses related to prostitution.[25]

[Insert table 4]

The key questions for the identification of peers in the paper are the following: "*Do you know other people who sleep on the street? If yes, how many?*" and "*Of these, could you please tell me the name and surname (or alternatively, the first three letters of the surname) of the first five friends on whom do you rely on in case of need?*". Approximately, 33% of the respondents knows more than 20 homeless individuals. Table 4 reports the distribution of homeless friends, by place

[24] The exact question was "*Have you ever been in prison?*". If yes, "*Have you been in prison before, after or before and after you slept on the street/shelter for the first time?*".

[25] In Italy, illegal immigrants apprehended by the police are not incarcerated. Indeed, the last reform on immigration policy introduced the possibility of incarceration for illegal immigrants but such norm was never enforced because it was deemed anticostitutional.

of interview.[26] [27] About 36% of homeless people do not rely on any homeless friends and this percentage is higher for people who slept in shelters during the night of the count. Each individual has an average of 1.35 friends. 21% has one friend and only about 6% reports names and surnames of five friends. Approximately, 11.4% of respondents refused to answer and this fraction is slightly higher among sheltered homeless. The second part of table 4 contains analogous figures for the sample of inmates. There is no ex-ante difference in the distribution of friends among inmates who have been in prison *before* being homeless and the one in the general sample (the test for the equality of the coefficients do not reject the null with a p-value of 0.621), while among those who went to prison *during* homelessness, we note a higher fraction of respondents with less friends (the test for the equality of the coefficients rejects the null with a p-value of 0.001). Incarceration rates are monotonically declining with the number of friends. Although this pattern does not necessarily represent the causal effect of network size on the probability of incarceration, it provides a first piece of evidence on the importance of friends in reducing the likelihood of imprisonment.

Another interesting point is to understand the length of a friendship and if two friends generally arrive on the street at the same time. The MHS includes information on the time of knowledge of each homeless friend. On average, respondents know their best friends since about 4 years. By looking at the correlation between respondent′s duration on the street and his peers' duration - extremely low and not statistically significant - it emerges that friends arrived on the street for the first time at a different date, suggesting that they are not stuck with the friends they met in the first instance (not displayed).

[Insert table 5]

Information on the name and surname of each respondent is also elicited. By matching friends' questionnaires with respondents' questionnaires, I am able to obtain information on the characteristics of the nominated friends. Names and surnames of respondents in shelters have been checked with administrative data provided by shelters' administration, while we consulted soup kitchens and

[26] To double check whether the names reported are actually related to homeless friends, I compared the period of knowledge of each friend with the respondent's duration on the street. The period of knowledge is always lower than respondent's duration, except for the case of two brothers and two spouses who nominated each other as friends.

[27] The survey did not elicit information on the name and the surname of non-homeless friends for two main reasons. First, the sociological literature describes the homeless as the most excluded people in society (Jencks, 1994) who have been abandoned by everyone, including friends, and have difficulties in building up pure relationships with people who do not live on the street (Anderson and Snow, 1993). Hence, I assume the fraction of non-homeless friends to be negligible. Second, analyzing relationships between homeless and non-homeless individuals would have implied a survey among non-homeless friends to elicit their criminal behaviors and socio-economic characteristics. This was beyond the scope of the paper.

social service centres' registers for unsheltered respondents' names.[28] Table 5 shows figures on one's peers criminal behavior. The first part of the table provides a breakdown of the share of peers who have been in prison *before* starting their homeless spell. I assume this share of peers to be equal to zero if the respondent does not have peers who have been in prison before becoming homeless or if he does not have any friends. 69.16% of homeless interviewed do not have peers with criminal records before their first homeless spell or do not have friends. 3.03% of the respondents have a peer group composed for one half by criminals, while more than 5% have only criminal friends. 19.25% cited friends not in the sample or who did not answer to the question on their best friends or on incarceration.[29] To sum up, about 11.6% of the sample have at least one peer who have been in prison before starting his homeless spell. The second part of table 5 reports the share of criminal peers for the sample of inmates, to check whether inmates have a higher propensity to meet more criminal friends. Inmates, both before and after being homeless, report, on average, a higher fraction of criminal peers compared to the total sample, suggesting the importance of tackling criminals' self-selection in the same network.

6 Empirical strategy

The goal of this section is to assess whether the size of social networks and peer characteristics have a role on one's probability to be imprisoned. As explained earlier, the likelihood of imprisonment is assumed to increase with the number of crimes committed and used as an imperfect proxy for criminal behavior.[30] I estimate the following specification:

$$P_i = \alpha + \beta N_i + \gamma E\left[P^{Before}|g_i\right] + zx_i + \rho_p + \varepsilon_i. \tag{2}$$

The dependent variable, P_i, is a dummy equal to one if the individual i went to prison at least once *during* his homeless window, and zero otherwise. N_i describes the total number of $i's$ best friends. $E\left[P^{Before}|g_i\right]$ is the fraction of $i's$ peers who went to prison *before* being homeless, where g_i

[28] For example, to find respondent's missing surnames (a small fraction of them declares only his/her name), I crossed information on the name, age, nationality provided in the questionnaire with name, *surname*, age and nationality coming from administrative data.

[29] The share of i's criminal peers is computed by cross-reference the nominated friends in the sample. For example, if the respondent cited the name of 4 friends but only 3 were in the sample and, among these, one individual went to prison before homelessness, the fraction of i's criminal peers is equal to $\frac{1}{3}$. A potential concern arises: criminals might be less likely to participate in the survey and thus, to be in the sample. In the above example, if the non-observed friend would have been a criminal, the share of i's criminal peers would have been equal to $\frac{2}{4}$. This is an inescapable limitation, but it implies an harmless downward approximation of the average number of criminals in the sample.

[30] As noted in Cameroon and Trivedi (2005), measurement errors in the dependent variable may inflate the standard errors but do not lead to inconsistent estimates.

represents $i's$ network and it includes all the nominated best friends. In this framework, peers are individual-specific. x_i is a vector of individual's traits, including age, age squared, a second order polynomial in the length of the homeless spell, gender, education, nationality, as well as a dummy variable equal to one if the individual i has already been in prison *before* becoming homeless. The latter controls for the individual propensity to be criminal independently by the homeless status and it guarantees that the effect of the covariates is studied on individuals with no prior homelessness criminal experience. ρ_p denotes place of interview fixed effects to adjust estimates for common unobservable shocks affecting people who reside in the same shelter or in the same street. Peer effects are measured by the parameters β and γ. Specifically, β captures the effect of the network size on subsequent criminal behavior, while γ measures the impact of the fraction of criminal friends.[31] The conjecture is $\beta < 0$ and $\gamma > 0$. These predictions will be tested in table 7 below, and subjected to a series of robustness tests.

In interpreting γ as social interaction effect, we are implicitly assuming that a potential peer influences i's criminal behavior because his imprisonment happened *before* the one of the respondent. Figure A.1 in the Appendix shows all the possible scenarios for the existence of peer exposure and the plausibly of this assumption is motivated and discussed in Section 7.

The estimation of equation (2) deserves further discussion. First, the number of friends, N_i, is not exogenous, either due to omitted variable bias and to reverse causality. Some unobservables affecting the probability of imprisonment might also affect the number of friends. For example, more self-confident individuals could be more involved in criminal activities, but they also might be more likely to have more friends: in this case, the size of the social network will be almost certainly positive correlated to the error term ε_i and conventional OLS estimates of the parameter β, will be biased upward relative to the true causal effect. Or, alternatively, individuals arrested during their homeless window could have a larger social network size because they met up friends in prison. In order to correctly identify the causal impact of network size on imprisonment, it is necessary to identify an instrument that is correlated with network size but uncorrelated to the error term, ε_i. I exploit the variation in rainfall shocks during one's period as homeless to instrument N_i as follows:

$$N_i = \alpha + \theta Rain_i + \eta Released_i + zx_i + \rho_p + u_i. \tag{3}$$

[31] Equation (2) assumes that peer effects operate through the influence of peer characteristics (or exogenous effects as defined by Manski (1993)), such as criminal history, rather than peer behavior (or endogenous effects). Indeed, in the context of criminal behaviors, the role of exogenous peer effects is easier to estimates and nevertheless crucial for policy implications.

$Rain_i$ is the fraction of rainy days during i's period on the street. Imagine that i arrived for the first time on the street on July 13^{th}, 2006: $Rain_i$ is the ratio between the sum of the rainy days from July 13^{th}, 2006 and January 15^{th}, 2008, the date of the survey, divided by the total number of days between July 13^{th}, 2006 and January 15^{th}, 2008. The idea behind the instrument is that homeless people are more likely to be in sheltered places during rainy days, such as bridges, porches or recreational centers, and they have a higher likelihood to meet more people and to make more friends. Hence, the prediction is $\theta > 0$. The variation in the $Rain_i$ variable comes from the different date (day/month/year) at which a given individual arrived on the street. The date of arrival is individual specific and it depends on economic and family related shocks during a person's life, which is plausibly as good as random. Those who arrived on the street for the first time in the same day, month and year have the same fraction of rainy days. Before explaining the variable $Released_i$, let's analyze potential concerns that could threat the validity of the instrument. The key identifying assumption is that, conditional on other covariates, the fraction of rainy days in one's period as homeless cannot directly influence the probability of imprisonment during homelessness, but it is also not correlated with any other (omitted) factors that could affect P_i. More formally, $Cov\left(Rain_i; \varepsilon_i | x_i\right) = 0$.

On this respect, previous studies in criminology have examined the correlation between weather and crime. This relationship appears to greatly vary with the weather condition (i.e. temperature, rain, wind and humidity) and the type of crime examined (DeFronzo, 1979; Perry and Simpson, 1987; Cohn, 1990). For example, assaults, burglary, collective violence, domestic violence, and rape tend to increase with ambient temperature, at least up to about 85°F. High temperatures do not appear to be correlated with homicide, robbery, larceny, and motor vehicle theft. It is not currently possible to draw any firm conclusion about the relationship between cold temperatures and crime, rainfall and crime, or wind and crime. In particular, the research on rain and crime suggests that rainfall does not significantly influence homicide and rape, and the relationship between rain and property crime is not clear. Although these studies do not allow for any unambiguous inferences about cause and effect, they seem to suggest that high temperature, more than rainfall variability, might influence the individual decision to commit a crime. In line with this hypothesis, in the economic literature, Jacob, Lefgren and Moretti (2007) examine the short-run dynamics of criminal behavior by using weather shocks, namely temperature and rainfall, to instrument the level of crimes in a jurisdiction: their findings show that while temperature is strongly correlated with property crime, the coefficient on precipitation is not statistically significant.

Furthermore, in the setting of this paper and given the most common types of crimes committed by the homeless (i.e. property crimes), it is hard to think about any direct correlation between an homeless's decision to perform illegal offenses and rainfall. For instance, if an homeless person desperately needs money to buy food, he probably would commit an illegal act with and without rainfall. On the other hand, rainfall would surely influence his decision to look for a sheltered place, resulting in greater opportunities for social interactions.

To further rule out other potential concerns about the validity of the exclusion restriction, I exploit two additional sources of data. First, I study the association between the monthly number of individuals arrested and the average monthly number of rainy days in Milan from 1998 to 2007: figures indicate a fairly low (.12) and not statistically significant pairwise correlation (p-value equal to .17). This correlation remains not statistically significant also controlling for year dummies. Second, I look at the crime reported by the police to the judiciary authority from 1983 to 2003 in Milan. These statistics are published yearly by the Italian National Institute of Statistics (ISTAT), and they allow a disaggregation by provinces and by type of crime. The correlation between the number of rainy days in each year and the yearly number of crimes reported in Milan is negative and equal to -.28, but, once again, not statistically significant (p-value equal to .21). With the same source of data, I focus on a series of crime categories that are most common among the homeless: property crimes (burglary, robbery, common theft), drug-related crimes and offenses related to prostitution: the p-value, equal to .16, confirms a non-statistically significant correlation between the number of rainy days and the number of crimes most commonly committed by homeless people.

A second concern related to the instrument validity is that rainfall may influence the intensity of non-criminal activities, and therefore indirectly affects crime. For example, if rainfall affects income for those who use to beg on the street, the criminal behavior of individual i may also be affected.[32] In table 2, we report evidence that none of the respondents declare begging as the first source of income and only 0.21% declare begging as second source. Hence, the results do not change by estimating equation (2) excluding individuals who use to beg.

Taken together, these pieces of evidence indicates that rainfall is unlikely to be directly correlated with the probability of imprisonment or the error term and it is a valid candidate to instrument the network size.

[Insert figure 3]

[32] For example, there could be a negative relationship between income and rainfall if during rainy days the homeless earn less money from begging due to less pedestrian traffic, or a positive correlation if during rainy days the homeless earn more money because they beg in shopping malls or in the underground.

Another natural concern regarding the use of past values of rainfall as instrument is whether there is sufficient variation in the rainfall to identify peer effects precisely. In other words, individuals who entered in the homeless status many years ago might have very similar rainfall distributions, while people who have not been homeless very long should be exposed to higher rainfall variation. In figure 3, I plot the fraction of rainy days, by respondent's duration on the street (in days). As expected, the graph shows a higher variation in the instrument for individuals with shorter homelessness spells. The spike in the fraction of rainy days for those who arrived on the street one or two months prior to the survey captures the high quantity of rainfall during the winter months before January, the date of the survey.[33]

A second problem in estimating equation (2) with a standard OLS model concerns the potential sorting of past criminals and potential criminals in the same network. Individuals with a higher propensity to commit illegal offenses could be more likely to look for homeless friends with previous criminal experience. To deal with this issue, I exploit the fraction of inmates released by Milan's authorities to instrument the share of criminal peers as follows:

$$E\left[P^{Before}|g_i\right] = \alpha + \eta Released_i + \theta Rain_i + zx_i + \rho_p + u_i \tag{4}$$

where $Released_i$ represents the fraction of inmates released from correctional facilities in Milan during one's period as homeless. More precisely, it is computed by dividing the number of released inmates since person i first became homeless and the number of months person i has been homeless.[34] The variation in the instrument is reported in Figure 3. Those who arrived on the street for the first time in the same month and year have the same fraction of inmates released. The idea behind the instrument is that exogenous government policies driving inmates' outflow increase the supply of criminal potential friends and, consequently, positively affect the likelihood to meet more criminal peers. I rely on the assumption that the share of inmates released in a homeless spell is not directly correlated with the probability to be imprisoned and it is also not correlated with any other factors that could affect P_i, that is $Cov(Released_i; \varepsilon_i) = 0$.

A concern related to the above assumption is that the share of inmates released may alter the costs and the benefits of criminal activities. If an increase in the number of inmates released correspond

[33] In particular, the maximum value for the fraction of rainy days (.88) is recorded by the homeless who spent 6 days on the street. This is computed by dividing the total number of rainy days, 5.33 days (i.e. 2 days of rainfall recoded in 6 weather stations - 2 days - plus 4 days of rainfall recorded in 5 out of 6 weather stations - 3.33 days) and the total number of days spent on the street.

[34] As described in section 4, data on inmates released are available only from 1993 onwards. Hence, to the individuals who became homeless before 1993 (9.7%), I impute the fraction of inmates released from January 1993 onwards.

to an increase in the number of policemen or police expenditures, the likelihood of incarceration in the structural equation (2) might be affected. To address this concern, I directly test for whether increases in the number of inmates released are associated with the amount of police employed in Milan. I find little evidence in support of this argument: the correlation between yearly data on police enforcement in Milan from 1993 to 2000 (Barbaglio, 2000) with the number of inmates released from Milan's correctional facilities is equal to -.40 and reports a p-value of .28, thus increasing confidence on the hypothesis that $Released_i$ is exogenous to P_i. In table 6 below (columns (1)-(4)) I also show that the share of inmates released does not affect the network size, reducing concerns that the number of released might influence the outcome variable not only through the share of criminal friends (α), but also through the network size (N).

Finally, an important data-related issue arises because we only observed individuals who were homeless in a specific date, when the survey was conducted. We do not observed criminal homeless who were in jail during the survey. Our estimates are based on a *stock* of individuals who were homeless on a particular date. This implies that the probability of being in the sample depends on the time spent in the state, and it is therefore higher for individuals with longer homeless windows (*length-biased sampling*) (Salant, 1977; Cameroon, Trivedi, 2005). High frequency criminal individuals might be imprisoned more often, and, on average, might have smaller length of homeless spells. As a consequence, they might be less likely to be represented in the stock population. Hence, our estimates are based on a sample of less criminal individuals compared to the true population, with a consequent downward bias compared to the true effect.

7 Results

7.1 Main results

The results reported in this section examine how the size of the social network and peer characteristics influence the probability of imprisonment for the individual i *during* his period as homeless. In all estimates, standard errors are robust and adjusted for clustering of the residuals at the place of interview level (shelters, parks, streets, stations, etc.). Summary statistics on the variables used in the regressions can be found in the Appendix Table A.1.

[Insert table 6]

Table 6 reports instrumental variable first stage estimates and reduced form results. I begin by analyzing the impact of rainfall on the probability of incarceration. According to the estimates of column (1), the fraction of rainy days during i's period on the street strongly and positively predicts the size of the social network. During rainy days, the concentration of homeless people in sheltered places (i.e. bridges, train stations, underground) is more likely to increase, with a consequent higher probability of social interactions. In term of magnitude, a one standard deviation increase in the fraction of rainy days during one's period as homeless is associated with having 0.27 more friends. The effect is statistically significant at 1% level. In column 2, I include the instrument for the share of criminal friends: while rainfall remains positive and statistically significant, $Released_i$ does not influence the network size. This finding is robust to the inclusion of a variety of controls, such as age, age squared, length of the homeless spell (in months), length squared (in months), gender, education, nationality and a dummy indicating the place of the interview (column (3)). Holding the other controls at the sample mean, one standard deviation increase in rainfall, increase the dependent variable by 0.33 friends.

The positive coefficient of the rainfall variable deserves further investigations. Column (4) report an alternative specification of the variable describing rainfall variation. In particular, I study the different effects of rainfall, based on its quartiles' distribution. I keep the same controls as in the baseline specification in column (3). It is interesting to note an increasing and monotonic effect of rainfall in predicting the number of friends. Being in the last quartile of the rainfall distribution positively affects the dependent variable and the coefficient is statistically significant at 1% level. In column (5), I further modify the rainfall variable by taking into consideration only the fraction of rainy days during one's first year on the street. This alternative specification takes into account a greater degree of variation in rainfall during a shorter period of time and a potential greater level of effort in searching homeless friends at the beginning of the homeless status. The fraction of rainy days is still positive and statistically significant, but the coefficient is slightly smaller in magnitude compared to the one estimated in column (3).

Columns (6) and (7) of table 6 report the first stage results for the share of peers who went to prison *before* being homeless on the fraction of inmates released by Milan's authority during $one's$ period as homeless. The mechanism behind this instrument is that exogenous policies driving inmates' outflow affect the supply of criminal potential friends, increasing the likelihood to meet homeless friends who already had previous criminal experience. According to the estimates in column (7), ceteris paribus, a one standard deviation increase in the fraction of criminals released during $i's$ homeless spell increases the the fraction of $i's$ peers who went to prison *before* being homeless by 3.9 percentage points. The Cragg-Donald test for weak instruments is above the critical value of 7.03

(at 10% bias toleration) and 4.58 (at 15% bias toleration), suggesting that rainfall variation and the fraction of inmates released could be considered valid instruments in this setting (Sock and Yogo, 2002).

In columns (8)-(9), I report the reduced form regressions. Looking at the reduced form equation can further mitigate concerns about weak instruments (Angrist and Kruger, 2001). In almost all cases, we note the expected signs of the coefficients of interest and a statistically significant correlation between the instruments and probability of incarceration.

[Insert table 7]

The main research question of the paper is to understand the role of peer characteristics and of the network size on individual criminal decision. Table 7 reports second stage estimates where the dependent variable is equal to one if homeless i has been arrested *during* his period as homeless. Columns (1)-(3) show OLS results. According to the baseline estimate in column (1), the likelihood to go to prison during homelessness decreases by 3.4 percentage points with one additional friend. A plausible interpretation of this finding is that friends represent a source of mutual insurance. Homeless with more friends have greater chances to survive on the street without committing crimes because their idiosyncratic shocks are shared among a higher number of individuals. An alternative interpretation of the negative coefficient on the network size could be that having more friends reduces the probability of arrest, conditional on crime. This would be the case if, for example, homeless with larger social networks would have a lower chance to be caught by the police - because protected by a higher number of friends - compared to those with smaller or no networks. While this reasoning could be true in some specific contexts (the Italian Mafia is a well-known example), it seems irrelevant among homeless people: they do not have credential to be trusted by the police and no incentive to be involved in a trial as witnesses to protect a friend. The estimated peer effects, γ, are captured by the share of i's friends who went to prison *before* becoming homeless. OLS estimates in column (2) reveal the presence of peer effects in crime: the share of peers who served previous jail sentence increases the likelihood that an individual, with no prior adjudication, will be arrested. The coefficient remains positive and highly significant with the inclusion of socio-demographic controls and place of interview dummies (column (3)). The OLS estimates are consistent with the hypothesis that the number of friends reduces the probability of imprisonment, while the share of criminal friends increases it. However, these estimates might reflect the effect of unobservables that might simultaneously impact the likelihood to have many friends and to be arrested.

Columns (4)-(6) contain the key results of this empirical section. They show IV estimates of the probability of incarceration, by exploiting the fraction of rainy day in one's period as homeless as excluded instrument for the size of the network and the variation in the fraction of inmates released for the share of one's criminal peers. By looking at the specification with network size only (column 4), estimates confirm a negative and statistically significant correlation between the number of friends and the likelihood of arrest. This result is robust to the inclusion of the variable describing the share of criminal peers (column (5))[35]. In term of magnitude, one additional friend leads to a 16.2 percentage points decrease in the probability of going to prison during an homeless spell. Controlling for the endogeneity of γ, estimates clearly indicate the presence of significant peer effects in the probability of committing criminal activities: the exposure to a greater fraction of peers who served previous jail sentence strongly increases the likelihood of committing illegal acts for an individual with no prior criminal experience. A one standard deviation increase in the fraction of peers who went to prison before becoming homeless increases the probability that homeless i will go to prison by approximately 20 percentage points. The results hold also by including additional socio-demographic characteristics. Recall that we include the dummy $Prison_i^{Before}$ to gauge peer effects only on those individuals who did not have prior criminal experience. The inclusion of $Prison_i^{Before}$, length of the homeless spell and length squared among the controls is important to increase the precision in the estimation of β and γ. However, the estimated coefficients on these variables may suffer from endogeneity bias. For this reason they should not be given a causal interpretation. The purpose of including them among the controls is purely to test whether the correlation between incarceration and the number of friends or the share of criminal peers is driven by prior criminal experience or by the length of the homeless window. Note that by including a dummy for Italians, I attempt to control for a potential different cost of committing illegal acts for Italians and immigrants (i.e. immigrants might face a higher cost of crime because it involves deportation).

[Insert table 8]

The main results reported in table 7 use only a single peer measure - the fraction of criminal peers in the group. This naturally leads to the question of whether other measures of criminal peers would reinforce or threaten peer effects. In table 8, I use as alternative endogenous variable: a dummy variable equal to one if the respondent reports the name of at least one criminal peer.

[35] The results are robust to the alternative specification of rainfall variable used to estimates the first stage in table 6. When using the fraction of rainy days in the first year as homeless as excluded instrument, the coefficient (standard error) on the network size is equal to -0.101 (0.06).

According to both OLS and IV estimates in columns (1) and (3), having at least one friend with prior criminal experience still increases the probability of one's incarceration, without affecting the sign of the coefficient on network size. Looking at the magnitude of the coefficient in column 3, a standard deviation increase in the one-criminal peer variable leads to a 21 percentage points increase in the likelihood of arrest. The effect is very similar to the one estimated in the previous table. In columns (2) and (4), I prove the robustness of OLS and IV estimates to a specification with controls. The coefficients β and γ maintain the same sign and remain statistically significant at 1% level.

Overall, the results in this section show that homeless individuals with many friends are less likely to be arrested, presumably because of the presence of mutual insurance mechanisms, while having friends with prior criminal experience increases the likelihood of imprisonment, probably due to an information channel in acquiring information and criminal skills.

7.2 Interpretation

To have more insights on the relevance and the potential policy implications of these results, I compute a simple exercise by comparing an increase in the network size by one extra *criminal* friend versus one extra *non-criminal* friend, on the probability of imprisonment, given the average network size and the average fraction of peers in prison *before* becoming homeless in the sample. The number of criminal friends, $N_{i,CR}$ is defined in the model as a fraction α of the total number of friends (N_i), and it is equal to 0.14.[36] First, let's suppose to increase the network size by one additional *non-criminal friend*, from 1.37 friends to 2.37 friends. The fraction α of criminal peers will then decrease from 0.102 to 0.058.[37] The variation in the average share of criminal friends, with and without one extra non-criminal friend is then equal to -0.044. Based on the coefficients on the network size and on the fraction of criminal peers estimated in table 7, we can now computing the probability of arrest given an extra non-criminal friend. It turns out that, ceteris paribus, the probability of arrest decreases with an extra non-criminal friend by approximately 19.9 percentage points.[38]

Second, we can do the same exercise by calculating the effect on the probability to go to prison with one additional *criminal friend*. The network size increases again at 2.37 friends, but, in this case, the number of criminal friends also increases from 0.14 to 1.14. The fraction of criminal friends

[36] By taking the share of peers in prison *before* being homeless and the average value of the total number of friends, I find: $N_{i,CR} = \alpha N_i$ =0.10*1.37=0.14.

[37] The fraction of criminal peers is equal to: $\alpha = \frac{N_{i,CR}}{N_i} = \frac{0.14}{1.37}$=0.102. and equal to $\alpha = \frac{N_{i,CR}}{N_i} = \frac{0.14}{2.37}$ =0.058 with one additional non-criminal friend.

[38] This result is obtained by substituiting the coefficients on the network size, on the share of criminal peers and the variation in the network size and in the fraction of criminal peers with an extra non criminal friend in equation (2): P_i =1 (-0.162)+(-0.044) (0.854)=-0.199.

α is then equal to 0.48.[39] For the same reasoning, it turns out that an additional criminal friend increases the likelihood of imprisonment by about 24.7 percentage points.[40] This means that having one additional criminal friend increases the probability to commit crime more, in absolute term, than how a non-criminal friend decreases it. In reality, however, it is difficult to distinguish ex-ante between "criminal" and "non-criminal" types. To derive a more clear policy implications, it is useful to compute the expected value of the probability to go to prison, by using as proxy for the probability to be criminal, the average number of offenders in the sample (29.6%). In conclusion, the expected value of going to prison decreases by 7 percentage points with one additional friend, independently by his or her criminal records before homelessness.[41] This implies that boosting friendship would be an important policy to reduce criminal behavior among the homeless.

8 Robustness

This section provides some further evidence of the robustness of the main findings. First, through out the paper, I found a strong positive correlation between rainfall shocks and the number of friends. The interpretation of this effect is that homeless individuals are more likely to interact during rainy days because they are concentrated in sheltered places. Let's assume now that the number of rainy days have nothing to do with the network size and the effect that we have estimated is generated by other sources of variation. To test this competing explanation, following De Giorgi et al. (2010), I artificially construct a random variable indicating the number of friends, with the same distribution as the original network size variable. I expect to find no significant rainfall effect when the network size is artificially constructed. Column (1) in table A.2 in the Appendix shows that the fraction of rainy days does not affect the artificially constructed network size. Similarly, I generate a *placebo* share of criminal peers, by artificially and randomly assigning an hypothetical share of criminal peers to each individual, distributed as the original variable. Column (2) in table A.2 reports no statistically significant correlation between the fraction of inmates released and the random share of criminal peers. As a second experiment, in column (3), I report second stage estimates of equation (2), substituting the real network size and the real share of criminal peers with the artificially constructed ones. While the actual network size and the actual share of criminal friends respectively decreases and increases the probability of incarceration, as shown in table 7, the random allocation of criminal friends does not have any statistically significant effect on the dependent variable.

[39] The fraction of criminal peers is equal to $\alpha = \frac{N_{i,CR}}{N_i} = \frac{1.14}{2.37}$=0.48.

[40] The likelihood of imprisonment has been computed as follows: P_i=1(-0.162)+(0.48) (0.854)=0.247.

[41] The likelihood of imprisonment has been computed as follows: $P_i = (0.296)0.247 + (1 - 0.29)(-0.199) = -0.07$.

Second, the effect of peer characteristics tested in the paper relies on the assumption that a potential peer influences individual criminal behavior because his imprisonment happened *before* the one of the respondent. In other words, to influence individual criminal behavior, friends must be criminals before meeting individual i. Figure A.1 in the Appendix shows all the possible scenarios for the timing of imprisonment for individual i and a potential friend j. Peer exposure is clearly identified in scenarios A and B, where peer's imprisonment before becoming homeless happened before $i's$ incarceration, independently on the date of arrival on the street for i and j. In scenario C, i became homeless after j and, in principle, his incarceration should have happened before j's incarceration. But, since in equation (2) we are considering i's imprisonment only during his homelessness and not before, $i's$ period in prison cannot be, by construction, before j's period in prison. Scenarios C is therefore impossible. Scenario D could be the only case in which there will not be peer exposure because i's imprisonment happened before j's incarceration. However, since we are considering only homeless peers, if i met j before j has become homeless, this peer would not be in the sample. On the other hand, by instrumenting the share of criminal friends, I address the reverse causality problem if i met j after j has become homeless.

These examples provide further confidence that the results presented in the previous section are driven by peer exposure.

9 Conclusions

This paper analyzes the influence of peers during a period of homelessness on each other's subsequent criminal behavior. The paper makes two main contributions. First, the data used come from the first representative survey in Europe among the homeless, conducted by the author in January 2008 in Milan, Italy. While prior empirical literature, studying peer effects on individual decisions, identifies peers at aggregate level, this survey elicits information on friends' names which allows me to precisely map each respondent's network. Second, to control for the non-random matching of individuals to their peers, the paper applies an instrumental variable strategy by exploiting rainfall variation to instrument the size of the network in one's period as homeless, and variation in the inmates released by Milan's authorities to proxy the fraction of criminal friends in the network.

The results provide strong evidence of peer effects in the realm of criminality among homeless individuals. Specifically, the individual probability of being arrested, for individuals with no prior criminal experience, increases with the exposure to peers with a history of criminal records. However, the same probability decreases as the total number of homeless friends (criminal and non-criminal) increases. These findings are robust to different identification strategies.

While the empirical analysis do not attempt to explicitly distinguish between potential channels through which peers with criminal records affect the individual probability of imprisonment and why homeless individuals with more friends are less likely to go to prison, anecdotal evidence collected among homeless people might provide some insights on the underlining mechanisms. Criminal peers seem mainly to increase knowledge about how and where to commit crimes, thereby decreasing the perceived costs and increasing the potential returns from committing crimes. Contrarily, the total number of friends acts as a safety net for people living on the street. Homeless people are constantly hit by idiosyncratic shocks and they develop mechanisms of informal insurance with friends to smooth shocks and to guarantee their survival on the street.

These results have relevant implications from a policy perspective. The negative correlation between the size of social networks and crime may provide an incentive to policy makers to boost social interactions among the homeless society. For instance, drop-in day centers that offer social activities for homeless people such as board games, group conversations, movies, may be beneficial for stimulating social interactions and reducing criminal activities. On the other hand, the existence of peer effects on criminal behaviors suggests that any reduction in the criminal histories of peers would lead to further reductions in crime. Hence, programs targeting rehabilitation in prison and housing assistance after release might have beneficial spillover from the reduction of former inmates who might become homeless and, consequently, reduce crime propagation.

Could these results and policies be extended to other settings or population? Given the recent increasing trend in the homelessness rates, as one of the main consequence of the economic crisis, and the new profile of a homeless person, researchers should carefully reflect on the thin line between urban poor and homeless people. Certainly, they are the most vulnerable population in richer countries, characterized by a high degree of social interactions and a high share of former inmates.

References

[1] **Anderson L., Snow D. A**. 1993. "Down on Their Luck. A Study of Homeless Street People", University of California Press.

[2] **Anwar, S., Fang** H. 2006. "An Alternative Test of Racial Profiling in Motor Vehicle Searches: Theory and Evidence", *American Economic Review*, 96(1), 127-151.

[3] **Allgood S., Warren R. S.**, 2003. "The duration of homelessness from a national survey",

Journal of Housing economics 12, 273-290.

[4] **Angrist J., Kruger A.** 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments", *The Journal of Economic Perspectives,* 69-85.

[5] **Bandiera O., Rasul I.**, 2006. "Social Networks and Technology Adoption in Northern Mozambique", *Economic Journal* 116, 862-902.

[6] **Bandiera O., Barankay I., Rasul I**. 2010. "Social Incentives in the Workplace", *Review of Economic Studies*, 77(2), 417-59.

[7] **Barbagli M.**, 2000. "Data on the Italy's police forces 1814-2001", Barbagli's personal archive.

[8] **Bayer, P., Hjalmarsson R., Pozen D.**, 2009. "Building criminal capital behind bars: peer effects in juvenile corrections", *Quarterly Journal of Economics*, 124, 105-147.

[9] **Bayer, P., Ross S. L., Topa G.**, 2008. "Place of work and place of residence: informal hiring networks and labor market outcomes", *Journal of Political Economy,* Vol. 116 (6), 1150-1196.

[10] **Becker, G.**, 1968. "Crime and Punishment: An Economic Approach". *The Journal of Political Economy* 76, 169-217.

[11] **Beaman L. 2011**. "Social Networks and the Dynamics of Labor Market Outcomes: Evidence from refugees", *Review of Economic Studies,* forthcoming.

[12] **Bertrand M., Luttmer E. F. P., Mullainathan S**. 2000. "Network Effects and Welfare Cultures", *Quarterly Journal of Economics* 115(3), 1019-55.

[13] **Bobonis G., Finan F.** 2009. "Neighborhood Peer Effects in Secondary School Enrollment decisions", *Review of Economics and Statistics*, 91(4): 675-716.

[14] **Braga M., Corno L.** 2011. "Being homeless: Evidence from Italy", *Giornale degli Economisti e Annali di economia,* forthcoming.

[15] **Bramoullé Y., Kranton, R**. 2007. "Risk sharing across communities", *The American Economic Review, Papers and Proceedings,* 97(2),.70-74.

[16] **Brent B.,** 2007. "A Repeated Observation Approach for Estimating the Street Homeless Population", *Evaluation Review*, 31; 166.

[17] **Burt M.**, 1992. "Over the edge: The growth of homelessness in the 1980s", *Urban Institute Press*, Washington DC.

[18] **Calvò-Armengol A., Patacchini E., Zenou Y**. 2009. "Peer effect and Social networks in education", *Review of Economic Studies*, 76, 1239-1267.

[19] **Calvò-Armengol A., Zenou Y**., 2004. "Social network and crime decisions: the role of social structure in facilitating delinquent behavior", *International Economic Review,* 45, 939-958.

[20] **Cameroon A. C., Trivedi P. K**. 2005. "Microeconometrics. Methods and Applications", Cambridge University Press.

[21] **Case A., Katz L. F**., 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", *NBER Working Paper No.* W3705.

[22] **Cohn, E. G.** 1990. "Weather and Crime", *British Journal of Criminology*, 30(1), 51-64.

[23] **Conley T., Udry C**. 2010. "Learning About a New Technology: Pineapple in Ghana", *American Economic Review,* 100(1).

[24] **Cook P., Gross K. A.** 1996. "A selective review of the social contagion literature", Terry Stanford Institute Working paper, Duke university.

[25] **De Fronzo J**. 1984 "Climate and Crime: tests of an FBI assumption", Environment and Behavior, 16, 185-210.

[26] **De Giorgi G., Pellizzari M., Redaelli S**. 2010. "Be as careful of the books you read as of the company you keep: Evidence on Peer Effects in Educational Choices", *American Economic Journal: Applied Economics*, 2, 241-275.

[27] **De Weerdt J., Fafchamps M.** 2011. "Social Identity and the Formation of Health Insurance Network", *Journal of Development Studies*, forthcoming.

[28] **Dercon S., De Weerdt J.** 2006. "Risk Sharing networks and insurance against illness", *Journal of Development Economics* 81, 337-356 .

[29] **Fafchamps M., Gubert F.** 2006. "The Formation of Risk sharing networks", *Journal of Development Economics,* 83(2), 326-50.

[30] **Fafchamps M., Lund S.** 2003. "Risk Sharing Networks in Rural Philippines", *Journal of Development Economics,* 71, 261-87.

[31] **Fisher N., Turner S.W., Pugh R., Taylor C.** 1994. "Estimated numbers of homeless and homeless mentally ill people in north east Westminster by using capture-recapture analysis", *British Medical Journal*, 308, 27-30.

[32] **Glaeser E., Sacerdonte B., Scheinkman J**. 1996. "Crime and Social Interactions", *Quarterly Journal of Economics*, 111(2).

[33] **Godlonton S., Thornton R**. 2010. "Peer effects in learning HIV results", *Journal of Development Economics*, forthcoming.

[34] **Honig M., Filer R. K**. 1993. "Causes of Intercity variation in Homelessness", *American Economic Review*, 83(1),.248-255.

[35] **Jencks, C.**, 1994. "The Homeless". *Harvard University press.*

[36] **Jacob B., Lefgren L., Moretti E**. 2007. "The dynamics of criminal behavior: evidence from weather shock", *Journal of Human Resources,* 42(3).

[37] **Katz, L.F, Kling J.R., Leibman J**. 2001. "Moving to Opportunity in Boston: Early results from a randomized housing voucher experiment", *Quarterly Journal of Economics,* 116, 607-654.

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

[39] **Kling J.R., Ludwig J.** 2007. "Is crime contagious?", *Journal of Law and Economics*, 50 (3).

[40] **Lochner L., Moretti E**. 2004. "The effect of Education on Crime: Evidence from prison inmates, arrest and self-reports", *American Economic Review* 94(1).

[41] **Ludwig J., Duncan G., Hirschfield P**. 2001. "Urban Poverty and Juvenile Crime: Evidence from a randomized housing-mobility experiment", *Quarterly Journal of Economics*, 116, 655-680.

[42] **Manski, C.** 1993. "Identification of Endogenous Social Effects: The Reflection problem", *Review of Economic Studies*, 60, 531-542.

[43] **Mas A., Moretti E.** 2009. "Peers at work", *American Economic Review*, 99(1),.112-145.

[44] **Munshi K**., 2003. "Networks in the modern economy: Mexican migrants in the US labor market". *Quarterly Journal of Economics* 118, 549-599.

[45] **O'Flaherty, B.** 1996. "Making Room: The economics of homelessness", *Harvard University Press,* Cambridge, Massachusetts.

[46] **Patacchini E., Zenou Y**. 2012. "Juvenile Delinquency and Conformism", *Journal of Law, Economics & Organization*, 28, 1, forthcoming.

[47] **Perry J. D., Simpson,** M. E, 1978 "Violent cimes in a city: enviromental determinants" *Enviroment and Behavior*, 19:77-90.

[48] **Piliavin I., Wright, E. R. B., Mare, D. R., Westerfelt H. A.** 1994. "The dynamic of homelessness", *Institute for research on poverty, discussion paper,* No.1035.

[49] **Quigley J., Raphael S., Smoloensky E**. 2001. "Homeless in America, Homeless in California", *Review of Economics and Statistics,* 83(1), 37-51.

[50] **Reiss A. J.** 1988**.** "Co-Offending and criminal careers", in Crime and Justice, Morris N., Tonry M. (ed), vol.10, University of Chicago.

[51] **Rossi H. P.** 1991**.** "Down and Out in America: The Origins of Homelessness", University of Chicago Press.

[52] **Sacerdote B.** 2001. "Peer effects with random assignment: results from Dartmouth roomates", *Quarterly Journal of Economics*, 116, 681-704.

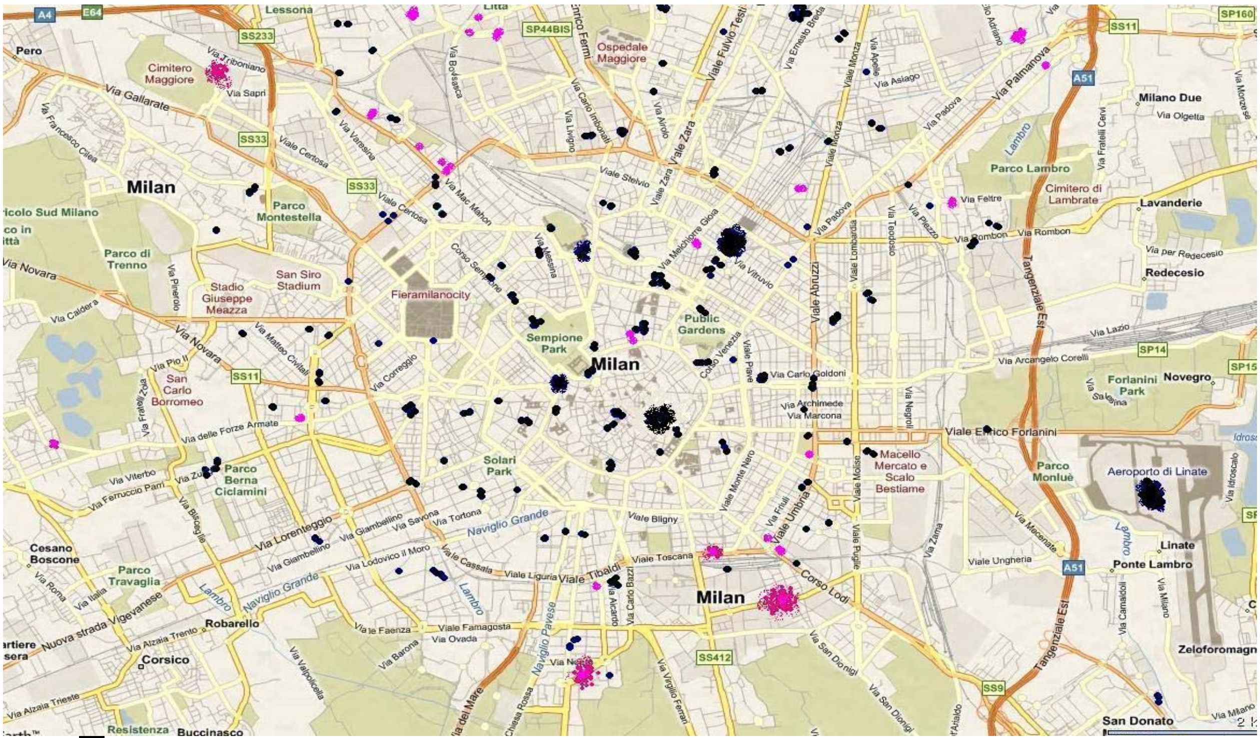
[53] **Salant S. W.** 1997. "Search theory and duration data: a theory of sorts", *Quarterly Journal of Economics*, 91, 39-57.

[54] **Stock J., Yogo M.** 2005. "Testing for Weak Instruments in Linear IV Regression", in Donald W. K. A. and James H. Stock, eds., Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, Cambridge: Cambridge University Press, 80–108.

[55] **Sutherland E. H.** 1947. "Principles of Criminology", Fourth Edition, Chicago: J.B. Lippincott.

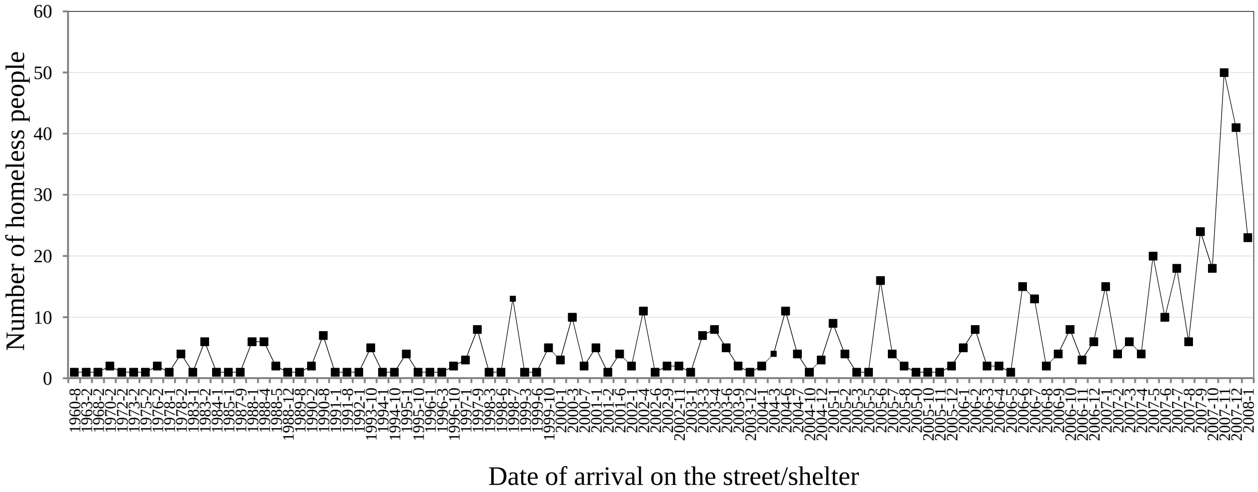
Figures

Figure 1: Spatial distribution of street homeless and shelters' locations. Milan, January 2008



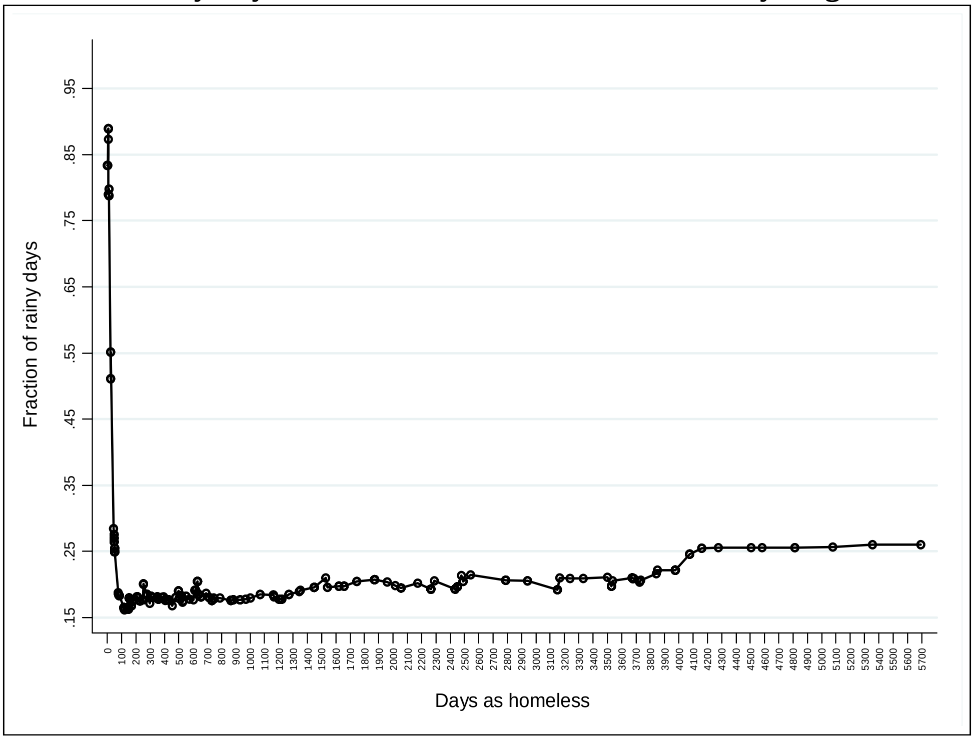
Legend: ■ = Location of street homeless: 1 dot equals to 1 homeless.
■ = Location of shelters: 1 dot equals to 10 homeless.
Source: MHS 2008.

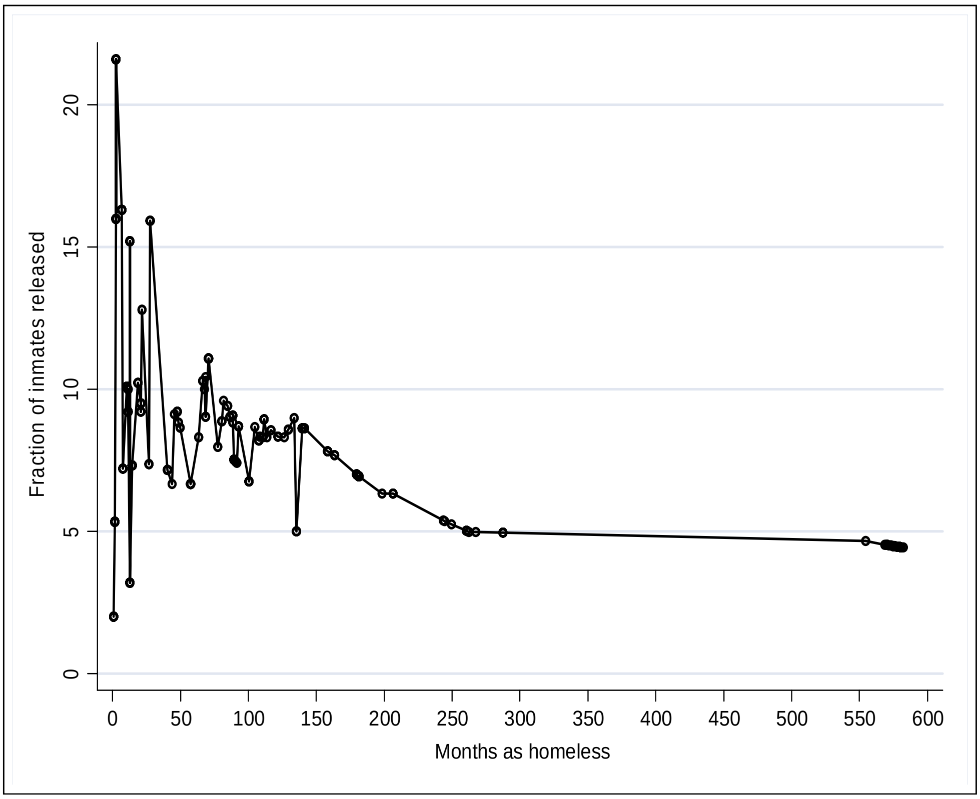
Figure 2: Date of first arrival on the street/shelter



Source: Author's calculations on the MHS 2008.

Figure 3: Fraction of rainy days and fraction of inmates released, by length of homelessness





Source: Regional Agency for the Environmental Protection and the Meteorological Department of the Military Aeronautics for rainfall data, Statistical Offices of correctional facilities in Milan for data on the fraction of inmates released.

Tables

Table 1: Descriptive statistics on the homeless count. Milan, January 2008

	Street	*Shelter*	*Slums*
N. of homeless counted	408	1152	2300
N. of homeless sampled	408	500	525
% of homeless			
Interviewed	34.60	84.00	66.50
Who refused	11.98	2.00	--
Not found	21.00	6.69	33.50
Not interviewed due to time constraint	0.00	7.30	--
Who were sleeping	16.40	--	--
Not interviewed because we did not send enumerators	16.00	--	--
With bad quality questionnaires	0.02	0.01	--
Observations	141	420	349

Source: MHS 2008

Table 2: Fraction of homeless who declared an income, by source

First source of income	*All sample*	*Street*	*Shelter*	*Italians*	*Immigrants*
No income	13.01	7.09	15	9.24	16.03
Government subsidies[a]	13.59	12.06	21.43	38.15	3.84
Permanent/Occasional work/Savings	31.90	26.24	34.51	25.71	37.82
Family/Relatives	5.53	5.67	5.48	4.02	6.73
Friends	5.53	8.51	4.52	2.81	7.69
Shelter subsidy/Church/Associations	2.32	2.13	2.38	2	2.56
Illegal activities	1.07	0.71	1.19	1.2	0.96
Don't know/Don't answer	20.31	37.59	15.47	16.87	24.36
Observations	561	141	420	249	312

Notes: (a) Government subsidies include welfare checks, unemployment benefits, disability insurance, pension. To be eligible for welfare checks it is required to be Italian and to be resident in Italy. Regular immigrants can benefit from disability/unemployment insurance and pension if they meet the eligibility requirements. The survey also investigates the second and the third source of income. As second source, 5.27% declares income from family and friends, 4.64% from work, 1.47% from government subsidies, 1.26% from illegal activities, 0.21% from begging, 0.21% from shelter subsidies, 0.84% church and voluntary organization. The 98.5% of the sample does not have a third source of income.

Table 3: Fraction of homeless who have been in prison, by nationality

	All sample		*Italians*		*Immigrants*	
	Obs.	*Percent*	*Obs.*	*Percent*	*Obs.*	*Percent*
Prison at least once	*162*	*29.62*	*95*	*38.15*	*67*	*21.47*
Prison only before homelessness	50	9.14	31	12.76	19	6.25
Prison before and during homelessness	10	1.83	6	2.47	4	1.32
Prison only during homelessness	102	18.65	58	23.87	44	14.47
Never in prison	*385*	*70.38*	*148*	*59.44*	*237*	*75.96*
Don't answer	*14*	*2.5*	*6*	*2.41*	*8*	*2.56*
Total	*561*	*100*	*249*	*100*	*312*	*100*

Source: Author's calculations on the MHS, 2008. Notes: Before and during homelessness means before and after the first night an individual slept on the street/shelter.

Table 4: Distribution of friends, by place of interview and incarceration (%)

Distribution of friends	*All Sample*	*Street*	*Shelter*	*Sample of Inmates*[a]	
				In prison before homelessness	*In prison during homelessness*
0 friends	*36.19*	*28.37*	*38.81*	*33.33*	*50.00*
At least 1 friend	*52.40*	*56.03*	*51.19*	*53.34*	*41.96*
1 friend	21.03	20.57	21.19	20.00	16.96
2 friends	13.19	16.31	12.14	11.67	10.71
3 friends	5.88	6.38	5.71	6.67	7.14
4 friends	6.24	7.09	5.95	10.00	5.36
5 friends	6.06	5.67	6.19	5.00	1.79
Don't know/ Don't answer	*11.41*	*15.60*	*10.00*	*13.33*	*8.04*
Mean	1.35	1.5	1.3	1.46	1.01
Observations	561	141	420	60	112

Notes: Author's calculations on the MHS 2008. (a) Inmates are those who declared to be arrested at least once in their life.

Table 5: Share of *i*'s peers who have been in prison before homelessness (%)

Share of peers in prison before homelessness[(a)]	*All Sample*	*Street*	*Shelter*	*Sample of Inmates*[a]	
				In prison before homelessness	*In prison during homelessness*
0 [(b)]	*69.16*	*64.54*	*70.71*	*63.33*	*63.39*
At least one peer in prison before homelessness	*11.59*	*14.89*	*10.48*	*21.67*	*25.9*
0.25	0.89	0.71	0.95	3.33	0.89
0.33	1.25	2.13	0.95	1.67	0.89
0.5	3.03	4.26	2.62	6.67	7.14
0.666	1.25	2.13	1.43	1.67	4.46
1	5.17	5.67	5.00	8.33	14.29
No peers interviewed/don't answer	*19.25*	*20.57*	*18.81*	*15.0*	*10.71*
Mean	0.10	0.121	0.093	0.15	0.21
Observations	561	141	420	60	112

Source: Author's calculation on the MHS 2008. Notes: (a) The share of *i's* peers is computed among the sample of homeless interviewed; (b) The share of *i*'s peers is equal to zero if the respondent has no peers who have been in prison before homeless or if he does not have any friend.

Table 6: The effect of rainfall on network size and of inmates released on the share of criminal peers

	First Stage							*Reduced Form*	
Dependent Variable=	*Network size (number of best friends)*					*Share of i's peers in prison before homelessness*		*1 if i has been in prison during homelessness*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Rain_i$	2.049***	2.218***	2.564***			0.121	0.130	-0.253***	-0.026
	[0.564]	[0.646]	[0.722]			[0.087]	[0.095]	[0.058]	[0.068]
$Released_i$		0.012	0.018	-0.005	0.012	0.008***	0.008***	0.005	0.016***
		[0.011]	[0.013]	[0.013]	[0.012]	[0.002]	[0.003]	[0.004]	[0.003]
$Rain_i$: 2° quartile				0.117					
				[0.177]					
$Rain_i$: 3° quartile				0.423**					
				[0.217]					
$Rain_i$: 4° quartile				0.493**					
				[0.178]					
$Rain_i$ during the first year					2.062***				
					[0.629]				
Controls[a]	no	no	yes	yes	yes	no	yes	no	yes
Place of interview dummy[b]	no	no	yes	yes	yes	no	yes	no	yes
R-sq.	0.48	0.48	0.49	0.48	0.48	0.03	0.06	0.01	0.09
Shea Partial R-sq.	0.03	0.03	0.03	0.02	0.01	0.02	0.03	--	--
F-test	13.19	6.94	6.93	3.65	5.88	8.03	4.54	--	--
Cragg-Donald test	17.30	14.29	9.17	8.20	9.69	14.29	9.17	--	--
Observations	535	535	535	535	535	535	535	535	535

Notes: * denotes significance at 10 percent level, ** at 5 percent level, *** at 1 percent level. Robust standard errors in parenthesis, adjusted for clustering at the place of the interview level (each shelter or street area). *Network size* is the total number of direct friends and it varies between 0 and 5. $Rain_i$ is the fraction of rainy days in one's period as homeless and it is computed as the total number of rainy days in one's period as homeless out of the total number of days as homeless. $Released_i$ is the fraction of inmates released by Milan's authorities during one's period as homeless. (a) Controls include a dummy equal 1 if homeless *i* went to prison before becoming homeless, age, age squared, length of the homeless spell (in months), length squared (in months), gender, education and a dummy equal to one if *i* is Italian. (b) A dummy equal to 1 if the respondent was interviewed on the street, 0 if in a shelter.

Table 7: Network size, criminal peers and the probability of imprisonment

Dependent Variable.: 1 if has been in prison during homelessness						
	OLS			*IV Estimates*		
	(1)	(2)	(3)	(4)	(5)	(6)
Network size	-0.034***	-0.050***	-0.050***	-0.159***	-0.162***	-0.127**
	[0.009]	[0.009]	[0.008]	[0.041]	[0.035]	[0.055]
Share of i's peers in Prison *before* homelessness		0.445***	0.484***		0.854*	2.262***
		[0.108]	[0.111]		[0.471]	[0.673]
$Prison_i^{Before}$			-0.107*			-0.213**
			[0.049]			[0.112]
Age			0.001			0.008
			[0.006]			[0.008]
Age sq.			0.000			-0.000
			[0.000]			[0.000]
Length of homeless spell (months)			0.002***			0.002***
			[0.001]			[0.000]
Lengh sq. (months)			-0.000**			-0.000***
			[0.000]			[0.000]
Female			-0.091*			-0.101
			[0.054]			[0.072]
Years of education			0.005			0.005
			[0.004]			[0.005]
Italian			0.053			0.021
			[0.048]			[0.045]
Place of interview dummy (a)	no	no	yes	no	no	yes
R-squared	0.02	0.09	0.16	0.03	0.13	0.17
F-test 1° stage I	---	---	---	13.19	6.94	6.93
F-test 1° stage II	---	---	---	---	8.03	4.54
Cragg-Donald test	---	---	---	17.30	14.29	9.17
Observations	547	547	535	535	535	535

Notes: * denotes significance at 10 percent level, ** at 5 percent level, *** at 1 percent level. Robust standard errors in parenthesis, adjusted for clustering at the place of the interview level (each shelter or street area). Network size is the total number of direct friends and it varies between 0 and 5. *$Rain_i$* is the fraction of rainy days in one's period as homeless and it is computed as the total number of rainy days in one's period as homeless out of the total number of days as homeless. *$Released_i$* is the fraction of inmates released by Milan's authorities during one's period as homeless. Controls include a dummy equal 1 if homeless i went to prison before becoming homeless, age, age squared, length of the homeless spell (in months), length squared (in months), gender, education and a dummy equal to one if i is Italian. (a) A dummy equal to 1 if the respondent was interviewed on the street, 0 if in a shelter.

Table 8: Network size, criminal peers and the probability of imprisonment: alternative measure of criminal peers

Dependent Variable.: 1 if has been in prison during homelessness				
	OLS		*IV Estimates*	
	(1)	(2)	(3)	(4)
Network size	-0.057***	-0.058***	-0.173***	-0.112***
	[0.009]	[0.008]	[0.032]	[0.043]
At least one peer in Prison *before* homelessness	0.343***	0.378***	0.689*	1.748***
	[0.068]	[0.074]	[0.397]	[0.448]
$Prison_i^{Before}$		-0.113**		-0.254***
		[0.052]		[0.104]
Age		0.0001		0.001
		[0.006]		[0.007]
Age sq.		0.000		0.000
		[0.000]		[0.000]
Length of homeless spell (month)		0.002***		0.002***
		[0.000]		[0.000]
Length sq. (months)		-0.000***		-0.000***
		[0.000]		[0.000]
Female		-0.091		-0.097
		[0.056]		[0.083]
Years of education		0.005		0.006
		[0.004]		[0.005]
Italian		0.045		-0.014
		[0.050]		[0.054]
Place of interview dummy [a]	no	yes	no	yes
R-squared	0.09	0.16	0.12	0.31
F-test 1° stage I	---	---	7.32	5.50
F-test 1° stage II	---	---	11.70	5.94
Cragg-Donald test	---	---	16.47	9.12
Observations	547	535	535	535

Notes: * denotes significance at 10 percent level, ** at 5 percent level, *** at 1 percent level. Robust standard errors in parenthesis, adjusted for clustering of the residuals at the place of the interview (name of the shelter/street). Constant not displayed. *Network Size* is the total number of direct friends and it varies between 0 and 5. *Prison Before*$_i$ is a dummy variable equal to one if homeless i has been in prison before the beginning of his homeless spell. (a) A dummy equal to one if the respondent was interviewed on the street, 0 if in a shelter. Columns 1-2 report OLS estimates, columns 3-4 show instrumental variable coefficients. Excluded instrument for the network size is the fraction of rainy days in *i*'s homeless spell. Excluded instrument for having at least one criminal friend is the fraction of inmates released during one's period as homeless.

Appendix

Figure A.1: Peer exposure

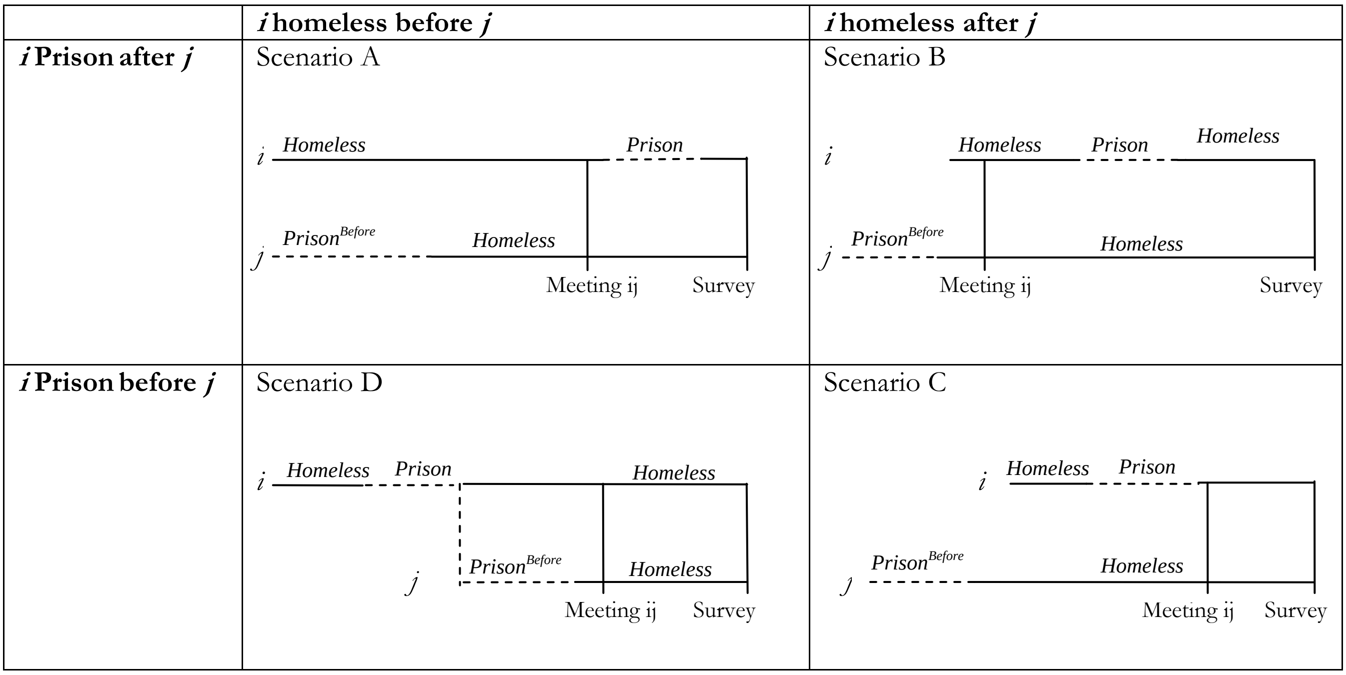


Table A.1: Descriptive statistics

Variable	***Obs.***	***Mean***	***Std. Dev.***	***Min***	***Max***
Prison$_i$	547	0.204	0.403	0	1
Network Size	547	1.37	1.47	0	5
Share of *i*'s peer in prison *before* homelessness [(a)]	547	0.10	0.24	0	1
At least 1 peer in prison	547	0.14	0.31	0	1
Rain$_i$	535	0.22	0.13	0.16	0.88
Rain$_i$ during the first year	535	0.13	0.16	0.00	0.88
Released$_i$	535	8.87	4.89	1.56	21.06
Prison$_i^{Before}$	547	0.10	0.31	0	1
Female	547	0.142	0.349	0	1
Age	547	44.68	13.35	19	82
Age sq.	547	2174.7	1233.1	361	6724
Years of educ.	547	9.12	4.08	0	20
Length of homeless spell (months)	535	107.80	167.75	0.5	582.5
Length of homeless spell sq. (months)	535	39710.6	97788.35	0.25	339306.3

Source: Author's calculation on the MHS 2008. Notes: *Rain$_i$* is the fraction of rainy days in one's period as homeless and it is computed as the total number of rainy days in one's period as homeless out of the total number of days as homeless. *Released$_i$* is the fraction of inmates released by Milan's authorities during one's period as homeless.

Table A.2: Robustness checks: random network size and random share of criminal peers

	First Stage		*Second Stage*
Dependent Variable:	*Random Network size*	*Random Share of i's peers in prison before homelessness*	*1 if i has been in prison during homelessness*
	(1)	(2)	(3)
Fraction of rainy days$_i$	0.188	-0.001	
	[0.664]	[0.096]	
Fraction of inmates released during i's homeless spell	-0.0323	-0.001	
	[0.015]	[0.002]	
Random network size			-0.137
			[0.008]
Random share of i's peers in prison *before* homelessness			0.056
			[0.062]
Controls	yes	yes	yes
Place of interview dummy[a]	yes	yes	yes
R-sq.	0.03	0.007	0.07
Shea Partial R-sq.	0.003	0.002	--
F-test	2.63	1.78	--
Cragg-Donald test	1.36	0.94	--
Observations	535	535	535

Notes: * denotes significance at 10 percent level, ** at 5 percent level, *** at 1 percent level. Robust standard errors in parenthesis, adjusted for clustering at the place of the interview level. Controls include a dummy equal 1 if homeless *i* went to prison before he became homeless, age, age squared, duration of an homeless spell (in months), duration squared (in months), gender, education and a dummy equal to one if *i* is Italian, month of entry in the homeless window. (a) A dummy equal to 1 if the respondent was interviewed on the street, 0 if in a shelter.