Homelessness and crime: Do your friends matter?

Lucia Corno*

July 10, 2015

Abstract

This paper investigates the influence of friends on crime, using data I collected among the homeless. To estimate the causal effects of friends and of the share of criminal friends on crime, I rely on two instruments. The first is the share of rainy days during one's first year as homeless: rainfall fosters homeless' concentration in sheltered places and increases the probability of interactions. The second is the share of inmates released during one's first year as homeless, which affects the supply of criminal friends. I find that one additional friend decreases the probability of incarceration but criminal friends increases it.

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

*Queen Mary, University of London and Institute of Fiscal Studies (IFS), Francis Bancroft Building, Mile End Road, London, E1 4NS. Email: l.corno@qmul.ac.uk. This project required the collaboration of many people. First of all, I am deeply indebted to Michela Braga, my co-worker during all stages of the project. The data collection benefitted 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. Thanks to Kjell G. Salvanes and two anonymous referees for valuable suggestions. I also benefitted from inputs from Richard Blundell, Stefano Dellavigna, Günther Fink, Randi Hjalmarsson, Magne Mogstad, Masayuri Kudamatsu, Luigi Minale, Enrico Moretti, Steve Pischke, Imran Rasul, David Strömberg, and seminar participants at Berkeley, Bocconi University, University College London, and IIES at Stockholm University. I thank the Regional Agency for Environmental Protection and the Meteorological Service of the Military Aereonautics for making rainfall data available to me, and the correctional facilities of San Vittore, Bollate, and Opera for providing me administrative data on inmates. 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, a homeless person in Milan

Homelessness is a critical problem among the urban poor and a major policy concern in many industrialized countries.[1] Because the homeless often live in extreme poverty, the expected benefits of crime are likely to outweigh the cost of potential punishment, thus making criminal activities more attractive (Becker, 1968). However, the lack of reliable and comprehensive surveys does not provide any guidance to policy makers in designing policies to reduce the crime rate among the homeless.

The goal of this paper is to investigate the influence of the size of the friendship network and of friends' characteristics on the criminal behaviour of the homeless, using original data I collected by interviewing homeless people in Italy. The previous literature indeed suggests that social interactions are particularly important in the criminal sector, where informal networks compensate for the lack of formal institutions in gaining knowledge and criminal skills.[2] Furthermore, because of their social exclusion and isolation, homeless peers may play a crucial role in shaping individual behaviour. Among the homeless, peers are the main source of information about potential jobs, shelter locations, and welfare programs, and they may also provide informal insurance against idiosyncratic shocks.[3] However, well-known identification issues—mainly endogenous group formation and reflection—make peer effects difficult to estimate.[4]

The novelty of the present paper is three-fold. First, the data come from the first representative survey in Europe among the homeless, conducted by the author in January 2008 in Milan, Italy. The survey is particularly suitable to study peer effects because it includes the names and surnames of each respondent and those of their five best homeless friends. Thus, I am able to precisely identify a close set of the individual's peer group. Information is also available on each respondent's criminal status before and after becoming homeless. The analysis is based on 883 homeless individuals (street,

[1] The Europe 2020 Strategy, the new socio-economic agenda lauched in 2010 to exit the economic crisis, has identified homelessness and housing deprivation as first-order priorities to be tackled by the European Member States (European Commission, 2010).

[2] The concept that criminal behavior is a *learned* behavior has been emphasized by Sutherland (1947) and, subsequently highlighted by, among others, Cook and Gross (1996); Glaeser, Sacerdote and Sheinkman (1996); Case and Katz (2001); Calvo-Armengol and Zenou (2004); Kling and Ludwig (2007); Bayer, Hjalmarsson and Pozen (2009); and Drago and Galbiati (2012).

[3] Randomized experiments in developing countries also show that peers have a considerable influence on the decisions of the poorest. 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 (to appear) 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.

[4] See Manski (1993), Moffitt (2001), De Giorgi, Pellizzari, and Redaelli (2010), and Black, Devereux, and Salvanes (2013), among others.

sheltered homeless, and slum dwellers) interviewed in one single night, with a high proportion of former inmates in the sample (26%). Second, the paper estimates the causal effects of the number of friends and of the share of criminal friends on (subsequent) criminal behaviour 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 first year as a homeless person, to instrument the total number of friends: rainfall fosters a concentration of homeless people in sheltered places (e.g. bridges, the underground, train stations) and increases the probability of social interactions. The second instrument is the number of inmates released per month by Milan's authorities during one's first year of homelessness. This is used to correct for the self-selection of past criminals and potential criminals in the same network. Finally, the lack of detailed data on friendship has so far limited the study of social interactions in terms of the effect of peer characteristics. This paper provides evidence of the influence of the size of one's network on criminal behaviour, thus suggesting policies to reduce crime.

The empirical analysis starts by investigating the effect of the total number of best friends and of criminal friends, proxied by the share of friends who have already been in prison before the current spell of homelessness, on the individual probability of incarceration during the current spell of homelessness. I consider different timings of imprisonment between the respondent and their friends (during and before being homeless) to disentangle the direction of causality from peers to the individual's behaviour. 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 first year of homelessness. Rainfall shocks randomly allocate individuals on the street, causing concentrations of homeless persons in sheltered places (i.e. bridges, the underground, train stations), which raises their probability of forming new friendships. I consider variation in the instrument within the first year of homelessness to capture a plausibly higher effort in searching for friends at the beginning of one's homeless status. Changes in the fraction of rainy days strongly and positively predict the number of friends. An increase by one standard deviation in the fraction of rainy days during one's first year of homelessness is associated with having approximately 0.5 more friends. The key identification assumption here is that the fraction of rainy days during one's first year on the street/shelter/slum does not directly influence the probability of imprisonment during homelessness. 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 one's social network with the number of inmates released per month by Milan's authorities during one's first year of homelessness. In this case, the underlying assumption is that exogenous policies driving inmates' outflow increase the supply of criminal potential friends, without directly affecting an individual's criminal behaviour. Ceteris paribus, an increase by one standard deviation in the share of former inmates increases the fraction of criminal friends by about 11 percentage points. Different timings in the beginning of each spell of homelessness provide individual variation in the instruments.

The first main result shows that the likelihood of incarceration during a spell of homelessness decreases by 12 percentage points on average with one additional friend. The interpretation of this finding is consistent with the idea that friends represent a source of insurance against shocks. Intuitively, the homeless with more friends have greater chances of surviving on the street without committing crimes, because their idiosyncratic and temporary shocks are shared among a higher number of individuals. Some attempts to test this potential mechanism have been made in the paper. In particular, I conduct a supplementary analysis to investigate the role played by (i) old friends; (ii) mutual friends; and (iii) a homeless person's entire social network (e.g. friends of friends) in their probability of incarceration. First, in long-term relationships, individuals have more chances of returning a favor in the future. So, if friends act as a safety net against economic shocks, the number of old friends should have a higher influence on one's criminal behaviour. behavioural evidence supports the idea that individuals cooperate better in infinitely repeated games (Coate and Ravaillon, 1993). Second, if shocks are correlated across closer friends, risk-sharing should be stronger in non-reciprocal relationships (Fafchamps and Gubert, 2007). In this respect, the empirical findings point to an underlying insurance motive behind social interactions. Finally, I test whether indirect friends (i.e. friends of friends) have an effect on criminal behaviour. The results show that it is not only the number of direct ties (at one degree of separation) but also the number of indirect ties (at higher degrees of separation) that influence future criminal behaviour. However, the strength of the effect of network size on one's probability of imprisonment decreases as the number of degree of separation increases. So, more individuals create a higher probability of insurance, but the results show that this effect diminishes in the degree of separation within one's friendship network. This finding is in line with many previous studies on social networks (see, for example, Fowler and Christakis, 2008).

The second main result indicates the presence of significant peer effects in the probability of committing criminal acts: an increase by one standard deviation in the share of peers who have served previous jail sentences increases the likelihood of incarceration for an individual with no prior criminal experience by 23 percentage points.

The rest of this paper is organized as follows. Section 2 provides a background on the data collected among homeless people and reviews the literature on social networks and crime. Section 3 describes the research design for the data collection and Section 4 illustrates the data and descriptive statistics. In Section 5, I present the identification strategy. Section 6 shows the main results and robustness checks and Section 7 concludes.

1 Background

1.1 Previous research on homelessness

Homelessness is a public policy issue in many countries. However, the lack of reliable and comprehensive surveys has limited economic research and effective strategies to prevent and reduce the phenomenon.

So far, only a few countries have included the number of homeless people in the official statistics of the population census or have developed *ad hoc* methodologies to 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 has managed counts of homeless people every two years on a national sample of 80 communities in different geographical areas. The HUD's most recent estimates indicate 610,042 persons in the US living in emergency shelters, transitional housing, or on the street on any given night (HUD, 2013). In 1990, for the first time, the US Census Bureau included a collection of data on homeless people in the decennial census in five cities (Chicago, Los Angeles, New Orleans, New York, and Phoenix). Up to now, official and systematic countrywide data have not been available for Europe. But, in a few countries, counts of the homeless population have been conducted by local agencies or NGOs. For example, in Italy, the last national count was produced in 2000 by the Zancan Foundation and the Commission on Social Exclusion of the Ministry of Social Affairs (Commissione, 2002). The final figure counts a population of 17,000 homeless people. As far as I know, the Milan Homeless Survey (MHS), conducted by the author in January 2008, is the first attempt to provide a comprehensive survey of the homeless population as well as a rigorous estimate of its size.

So far, previous research on homelessness has investigated the causes of the substantial increase in the incidence of homelessness during the 1980s in the US, using data provided by HUD (Honig and Filer, 1993; O' Flaherty, 1996; Quigley, Raphael and Smolenski, 2001, Mansur, Quigley, Raphael, Smolensky, 2002) or by the US Census Bureau (Burt, 1990). A common result of these studies is that variation in homelessness arises from changed circumstances in the housing market and in the distribution of income. Specifically, tougher housing markets are positively associated with higher levels of homelessness. Some other papers have looked at the determinants of the duration of spells of homelessness (Piliavin, Wright, Mare and Westerfelt, 1994). Homeless spells are generally longer for those with a history of drug or alcohol abuse, while having received government benefits in the past seems to have a controversial effect on the average length of a spell of homelessness (Allgood and Warren, 2003; Braga and Corno, 2012). The data at hand allows me to provide a unique picture of the characteristics of the homeless population and to study the role played by friends in shaping its criminal behaviour.

1.2 Previous research on social networks and crime

Starting with the seminal book by Sutherland (1947), criminological and economic studies have recognized the crucial role of peers in shaping criminal behaviour. Peers may influence criminal activities by transferring skills (Glaeser et al., 1996), by sharing information (Calvò-Armengol and Zenou, 2004) and by affecting the social stigma associated with illegal acts (Cook and Gross, 1996).

However, only a few papers have been convincing in measuring the causal impact of peers on individual criminal decisions. There are major difficulties in estimating the influence of social interactions on criminal behaviour. First, the scarcity of publicly available data at the individual level limits the identification of social network boundaries at the aggregate level. Several papers studying neighborhood effects on criminal behaviour have shown that living in a neighborhood with a high crime intensity significantly raises one's probability of becoming a delinquent (Case and Katz 1991; Ludwig, Duncan and Hirschfield 2001; Kling, Ludwig and Katz, 2005; Kling and Ludwig, 2007; Damm and Dustmann, 2014). More recent literature considers correctional facility boundaries to study peer effects on post-release behaviour (Bayer, Hjalmarsson and Pozen, 2009; Drago and Galbiati, 2012). However, being part of the same neighborhood or of the same correctional facility does not necessarily imply social interactions between these individuals.

In this respect, a great advantage of the data collected for the present paper is the possibility of delineating a close set of each individual's peer group, identified by actual friends' nomination.[5]

[5] Only a few surveys have recorded information on the individual's or household's social networks based on friends' nomination: the *National Longitudinal Study of Adolescent Health* (Patacchini and Zenou, 2012, Liu, Patacchini, Zenou, Lee, 2013); a panel household survey conducted in Tanzania (Dercon and De Weerdt, 2006; De Weerdt and Fafchamps, 2011); and survey data from the rural Philippines (Fafchamps and Lund, 2003; Fafchamps and Gubert, 2006).

The other difficulties in identifying the causal effect of social interactions are related to endogeneity, i.e. to the extent that unobservable characteristics affecting the likelihood of having 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 to reflection, since in a peer group everyone's behaviour affects the others' behaviour and it is problematic to disentangle an individual's behaviour from that 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 activity (Ludwig, Duncan and Hirschfield 2001; Kling, Ludwig and Katz, 2005; Kling and Ludwig, 2007). Looking at non-randomized identification strategies, Bayer et al., (2009) analyse the influence that juvenile offenders serving time in the same correctional facility have on each other's subsequent criminal behaviour. To control for the nonrandom assignment to facilities, they include facility and facility-by-prior-offense fixed effects, thereby estimating peer effects using only within-facility variation over time. They find that having peers with a history of committing a particular crime increases the probability that an individual who has already committed the same type of crime recidivates with that crime. Drago and Galbiati (2012), by exploiting the 2006 Italian prison pardon, called the Collective Clemency Bill, find that a former inmate's decision to re-offend is influenced by peers' residual sentences at the date of release.[6]

In this paper, I propose an instrumental variable approach, by relying on individual variations in the timing of the beginning of each spell of homelessness. Furthermore, prior empirical research mainly tests the effect of criminal peers on individual criminal behaviour, while the effect of the total number of friends has been so far overlooked.

2 Survey design and data

2.1 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 survey involved about 350 volunteers, recruited from service providers to the homeless (e.g. the Red Cross), but also from

[6] The Collective Clemency Bill provided an immediate three year reduction in detention for all inmates who had committed a crime before May 2, 2006. The bill stipulates that if a former inmate commits another crime within 5 years of their release from prison, they will be required to serve the residual sentence suspended by the pardon (varying from 1 to 36 months) in addition to the sentence for the new crime. The policy effectively transforms one month of an original sentence into an additional one month of sentence for future crimes committed at the individual level (Drago and Galbiati, 2012).

students and private citizens, thanks to the substantial attention received by the project from local media and newspapers.

A survey among homeless people faces 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.[7] The second challenge arises since homeless people are both territorially mobile and likely to enter into and exit out of the homeless state, and the risk of counting and interviewing the same person twice is therefore very high. The Milan Homeless Survey 2008 (MHS 2008) included two major phases: 1) a point-in-time count; and 2) a comprehensive survey via trained interviewers. The Point in time survey, or the S—Night approach (Shelter and Street Night) aims at identifying the number of homeless people sleeping on the street, in shelters, and in slums simultaneously in one reference night in the whole town. As such, it ensures a minimum of double counting.[8] 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 any homeless hidden from public view during late-night hours. We applied some efforts to overcome this problem 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. Enumerators were asked to walk every street and other public place in their assigned area. To reduce the risk of skipping some streets, we provided them 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 one homeless person, while abandoned car/caravan enumerators tried to distinguish 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 (e.g. torches, food, beverages, and notebooks). Besides counting, the volunteers had 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

[7] 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.

[8] Among the most recent approaches 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 over 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 (2012) provide a detailed overview of different methods used to count homeless populations.

the sleeping place (e.g. 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: key information for testing for a potential sample selection. Volunteers combined these statistical activities with hot beverages and food distribution.

In the meantime, other teams of enumerators 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 Roma) are generally monitored by the municipal police. During the three months prior to the survey, project leaders visited the slums, to identify the typology (authorized/unauthorized), the type of ethnic group, and the number of people living in each area. During the field visits, we asked for permission to interview people in the slums and we announced the date of the survey. On the night of the count, enumerators only checked the sizes and locations of pre-identified slums. The average duration of the count was about three hours, from 10 pm to 1 am.

The count was necessary to have a precise idea of the phenomenon's size and to construct a census from which to randomly select 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.[9] 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 a full census of the homeless counted by sending back enumerators to the locations identified during the count. Sheltered homeless were randomly sampled from the population on the basis of the shelter's size. We created a random sample proportional to the shelter's size by over-sampling small shelters and under-sampling big shelters. We agreed with each shelter's manager in advance on the best time to run the interviews. Out of 25 shelters, four refused to participate and one had no guests at the time of the survey. Some interviews were conducted directly by the shelter's managers. Finally, the slums were sampled through a stratified random sampling method based on geographic location, typology, and size. More specifically, we stratified them according to city administrative division (9 areas), official area classification (authorized/unauthorized, shacks, abandoned buildings), and the

[9] The interviews in the slums were done on Saturday afternoon. We decided not to go into the slums during the night for security reasons.

area's size. We selected a total of 12 out of 56 slums. Within each selected area, we randomly extracted the respondents. During the interviews, the volunteers also distributed sanitary napkins and kids' clothes to the households in the slums. To preserve the enumerators' safety, we informed the police about the initiative without directly involving them.

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 try to ensure that the homeless counted would be the same as those interviewed, we included as the first question in the survey "Did you sleep here last night?" and if not, "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 than the general population. To address this drawback, we asked the enumerators from the Red Cross to fill a one page questionnaire describing the respondent's condition at the end of the survey. We did not consider questionnaires conducted with drunk or mentally ill homeless persons. As an incentive, the 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.[10]

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, the enumerators minutely checked for the presence of homeless people by walking along all the streets in Milan: there would not have been 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 of observing more 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 in Milan at its lowest 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 the shelters are most full will probably 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 the 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).

[10] The questionnaire is available upon request.

2.2 Additional Data Sources

The present paper exploits additional data sources. First, I use rainfall data to proxy the size of the friendship network. The rainfall data came 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 the municipality of Milan, but some stations lack data for some days of the month.[11] I calculate the average number of rainy days among the weather stations as a proxy for the daily rainfall in Milan.[12]

Second, I assembled administrative data from the Statistical Offices of three correctional facilities in Milan (San Vittore, Opera, Bollate) on released inmates. Specifically, these data include the monthly number of released inmates from 1993 to 2008. To proxy the number of former inmates in Milan, I compute the monthly sum of released inmates among the three correctional facilities.

Finally, the Italian National Institute of Statistics (ISTAT, 2014) provided data on quarterly economic indicators (GDP in level at market price and the number of unemployed) in the region of Lombardy, as a proxy for the economic conditions in Milan (Milan is the capital of Lombardy).

3 Descriptive statistics

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

[Insert table 1]

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, 21% were not found, and 16.4% were already sleeping at the time of the interview. Due to time constraints, we did not send enumerators to 16% of the identified locations.[13] In the shelters, we randomly sampled 500 individuals out of the 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 the slums, we randomly selected a sample

[11] The six weather stations in Milan are located in Lambrate, Parco Nord, Zavattari, Confalonieri, Juvara, and Brera.

[12] As defined by the ARPA, a rainy day records at least one millimeter of rainfall over 24 hours.

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

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 of 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 total sample of 910 observations and a final sample of 883 observations with no missing information about friendship and imprisonment.

To have a better insight on the magnitude of any possible sample selection, I compare data on gender and age recorded during the homeless count with those of the homeless interviewed. The percentage of women interviewed is exactly the same as the percentage of women in the total homeless population (10%) and the percentage of homeless older than 35 years is very similar in the sample to that in the total population (72.1% versus 72.4 %). This comparison provides some confidence that the homeless sample we are analyzing is representative of the homeless population in Milan.[14]

[Insert figure 1]

Milan is the second-largest city in Italy and has a population of about 1.35 million. Figure 1 presents the spatial distribution of street homeless and the location of shelters and slums 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 the street homeless are almost equally spread over the city. As shown in Figure 1, in Milan there are 25 shelters, and approximately 30 slums, mainly located in the suburbs.

[Insert figure 2]

A first relevant question to answer is: what are the main reasons driving people to live on the street? Some of them, such as unemployment and poverty, can be predictable, but others are less intuitive. Figure 2 shows the main reason for homelessness reported by Italians and by immigrants. 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 27% of Italians, suggesting that family is an important source of insurance against economic and

[14] Unfortunately, there is no formal way to test a potential sample selection of non-criminal homeless since information on criminal behaviors has been elicited only through the questionnaire. However, we believe this is a minimal concern since the enumerators were Red Cross volunteers in uniform, a clear sign that the survey was totally unrelated to police activity.

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, and the labor and the housing markets. Culture and ethnicity is one of the main reason for an individual to live in a slum. Among the other reasons, 9.7% of Italians report drug or alcohol abuse, compared to 1.7% of immigrants. Only 7% of Italians and about 1% of immigrants reported previous conviction as the main reason for their homelessness.[15]

[Insert table 2]

Table 2 presents summary statistics for the main variables used in the present paper. Women constitute 28% of the sample. The average age is about 39 years, ranging from 19 to 83 years for street homeless, from 19 to 74 for sheltered homeless, and from 9 to 66 for slum dwellers. Among the homeless interviewed in Milan, 31% are Italians and 69% are immigrants. Homeless people from Romania make up the largest group (29%), followed by the homeless from North Africa, mainly Moroccan and Algerian (10%). More than half of the sample declared having a primary education certificate and 30% a secondary education diploma. Looking at the marital status of the homeless, we note that many of them are single (39%), 3% are widowed, 19% are divorced, and about 34% are currently married. These figures are in line with the literature arguing that the family may represent a source of insurance for their members against economic shocks (see for example, Bentolilla and Ichino, 2007).

The Milan Homeless Survey elicits information on homeless criminal status prior to becoming homeless and during the spell of homelessness. It asks: '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?'. Table 2 shows a strong relation between homelessness and crime: 26% of the homeless have been in prison at least once (34% of Italians and 25% of immigrants). Approximately 10% of the sample declares having had a period in prison before being homeless, and almost 21% have been in prison during their period of homelessness, showing how in extremely poor conditions, crime could become more frequent. In the empirical part, the probability of imprisonment during homelessness is the main outcome. Following Lochner and Moretti (2004), I use the likelihood of imprisonment

[15] The main reasons for homelessness are very similar to the statistics coming from the most recent survey conducted in San Francisco in 2007 (HUD, 2008), except for two main differences. The first one is related to the figures on disabilities and illness: about 34% of the San Francisco homeless cited disabilities as the main cause for their status, compared to 4% of those surveyed in Milan. The second one is about the inability to pay rent or mortgage, which is the third most common reason among the homeless in San Francisco.

as a proxy for criminal behaviour throughout the analysis, assuming that a person's probability of conviction is an increasing function of the number of offenses committed. A limitation of the survey is that it does not elicit information on when the crime was committed or on the type of criminal offense. To have some insights into the types of crimes generally committed by homeless people, I assembled administrative data from correctional facilities in Milan on criminal acts committed by 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.[16]

[Insert figure 3]

A key feature of the data is that each respondent reports a different date of arrival on the street/shelter/slum (day/month/year) and, consequently, a different length of their spell of homelessness, computed by taking the difference in days or months between the date of the first time an individual slept on the street/shelter/slum and the date of the survey.[17] Figure 3 presents the number, by date (month/year) of arrival on the street (or shelter or slum) of homeless people interviewed. The average duration of a spell of homelessness is about seven years, differing by type of place of interview: about seven years for individuals interviewed on the street, about four years for those living in emergency shelters, and about ten years for slum dwellers.

[Insert table 3]

The survey elicits detailed information on friendship. Each individual was asked to report the name and surname of their best friends through the following questions: 'Do you know other people who sleep on the street/shelters/slums?' 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 you rely in case of need?'. The names and surnames of respondents in shelters have been checked with administrative data provided by the shelters' administration, while we consulted soup kitchens and social service centres' registers for unsheltered respondents' names.[18] Table 3, panel A presents the distribution of friendship. In the sample, 62.51% of the respondents declare the name and surname

[16] 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 a norm was never enforced because it was deemed unconstitutional.

[17] In an ideal setting, I would use the date of exit from the status of homelessness to calculate the end of the spell of homelessness. Unfortunately, this was impossible to recruit formerly homeless people for the survey.

[18] For example, to find respondent's missing surnames (a small fraction of them declared only their first name), I cross-checked information on the name, age, and nationality provided in the questionnaire with the name, surname, age, and nationality coming from administrative data.

of at least one friend, with about 24.46% reporting the name and surname of one friend and 9.85% reporting the name and surname of five friends. The distribution of friends is very similar for people who slept in the street (column 2) to that of those sleeping in a shelter (column 3) during the night of the count, while slum dwellers report on average a lower number of friends. By matching the respondents' questionnaires with their friends' questionnaires, I have been able to obtain information on the characteristics of the nominated friends. Table 3, panel B, shows figures on the share of respondent's friends who have been in prison before starting their spell of homelessness. Note that we have fewer observations than those presented in panel A because the share of friends in prison before homelessness has been computed by dividing the number of friends who have been in prison before becoming homeless by the number of friends interviewed in the survey. It is therefore missing for respondents with no friends or who cited friends not in the sample. 28.27% of the homeless interviewed cited the name and surname of at least one friend with a criminal record. 5.74% of the respondents have a peer group at least half composed of criminals, and 17.4% have a share of criminal friends equal to unity.[19]

[Insert table 4]

Table 4 shows the different categories of friendship links in our sample of homeless people (Panel A) and for those with at least one best friend (Panel B). In columns 1 and 2, I split the sample between respondents with 'old' friends and those with 'new' friends. 'Old' friends are defined as nominated friends with a spell of homelessness greater than or equal to that of the respondent; 'New' friends are defined as nominated friends with a spell of homelessness shorter by at least one year than the respondent's spell. A higher proportion of respondents have at least one old friend (nearly 45% in the full sample and 72% in the sample of respondents with at least one friend) than have at least one new friend (23% in the full sample and 36.7% in the sample of respondents with at least one friend). Column 3 presents the share of mutual (or reciprocal) nominations (homeless people who list one another as friends), column 4 shows the fraction of nominators only (respondents who listed friends but were not listed as friends by these same homeless people) and column 5 presents the nominees only (those who were listed by someone but did not list the same nominator). In the

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

full sample, nearly 23% of the respondents have at least one mutual friend (column 2) and the same statistic is equal to 36.8% in the sample of the homeless with at least one friend (column 3). Among non-reciprocal relationships, 47.1% of the respondents (91.7% in the sample of the homeless with at least one friend) are nominators (column 4) and 36.5% (44.6% among respondents with at least one friend) are nominees (column 5).

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 a society (Jencks, 1994), who 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 best friends to be negligible. Second, to analyse the relationships between homeless and non-homeless individuals, we would have to have conducted an additional survey among the non-homeless friends to elicit their criminal behaviour and socio-economic characteristics. This would have been beyond the main scope of the homeless survey.[20]

4 Empirical strategy

The goal of this section is to assess whether the number of friends and friends' previous convictions have an effect on one's probability of being imprisoned. I estimate the following specification:

$$Prison_{ipm} = \alpha + \beta NFriends_{ipm} + \gamma ShareCriminalsFriends_{ipm} + zX_{ipm} + \rho_p + \varsigma_m + \varepsilon_{ipm}. \quad (1)$$

where $Prison_{i,p,m}$, is a dummy equal to one if individual i interviewed in place p who arrived on the street in month m went to prison at least once during their spell of homelessness, and zero otherwise. $NFriends$ is the total number of i's best friends (reciprocal and non reciprocal); $ShareCriminalsFriends_i$ is the fraction of i's friends who had already been in prison before being homeless. The idea is to capture the influence of ex-ante criminals on one's probability of imprisonment. X_{ipm} is a vector of the individual's traits, including gender, age, age squared, education, dummies for the three main nationalities in the sample, length of the spell of homelessness, a dummy for a spell of homelessness shorter than a year, and a binary variable equal to one if the individual i had already been in prison before becoming homeless. The latter controls for any individual

[20] To check that the names and surnamers of nominated friends are actually of some homeless friends, I compared the period of knowledge of each friend with the respondent's duration on the street/shelter/slum. The period of knowledge is always less than the respondent's duration, except for the case of two brothers and two spouses who nominated each other as friends.

propensity to be a criminal independently of the status of being homeless, and it guarantees that the effect of the covariates is studied on individuals with no prior criminal experience while homeless. ρ_p denotes place of interview fixed effects to adjust the estimates for common unobservable shocks affecting people who reside in shelters, in the street or in slums. Finally, ς_m are 12 dummies for the month of entry in the spell of homelessness, to capture any unobserved shock at the monthly level (e.g. colder versus warmer months). ε_{ipm} is the error term. In some specifications, I also add controls for (i) the status of the economy in Milan: the GDP in level (at market price) by quarter and the number of unemployed (those aged 15 years or more looking for a job) by quarter in the region of Lombardy (Italian National Institute for Statistics, 2014); and for friends' characteristics, such as their average numbers of years of education, their average duration of the spell of homelessness, and the share of friends with the same nationality as that of the respondent.

Our coefficients of interest are β and γ. β captures the effect of the total number of friends on subsequent criminal behaviour, while γ measures the impact of the fraction of criminal friends.[21] Equation (1) will be tested in Table 7 below and subjected to a series of robustness tests.

The estimation of Equation (1) deserves further discussion. First, the number of friends, $NFriends_{ipm}$, is not exogenous, either due to omitted variable bias or to reverse causality. Some unobservables affecting the probability of imprisonment may also affect the number of friends. For example, more self-confident individuals could be more involved in criminal activities, but they may also be more likely to have more friends: in this case, the size of the social network will be almost certainly positively correlated with the error term ε_{ipm}, and so conventional OLS estimates of the parameter β will be biased relative to the true causal effect. Alternatively, individuals arrested during their homeless window could have more friends because they made friends in prison. In order to correctly identify the causal impact of the total number of friends on imprisonment, it is necessary to identify an instrument that is correlated with the number of friends but does not have a direct impact on the outcome. I exploit the variation in rainfall shocks during one's first year as a homeless person to instrument $NFriends_{ipm}$ as follows:

$$NFriends_{ipm} = \alpha + \theta Rain_{ipm} + \eta FormerInmates_{ipm} + zX_{ipm} + \rho_p + \varsigma_m + u_{ipm}. \tag{2}$$

$Rain_{ipm}$ is the fraction of rainy days during i's first year as a homeless person. Imagine that i

[21] Equation (1) 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 behavior, the role of exogenous peer effects is easier to estimate and nevertheless crucial for policy implications.

arrived for the first time on the street/shelter/slum on July 13^{th}, 2006: $Rain_{ipm}$ is the ratio between the sum of the rainy days from July 13^{th}, 2006 to July 13^{th}, 2007, and the total number of days between July 13^{th}, 2006 and July 13^{th}, 2007 (i.e. 365 days for respondents with a duration of homelessness longer than one year or the number of days on the street/shelter/slum for respondents with a duration of their homelessness shorter than one year). The idea behind the instrument is that rainfall influences the decision of a homeless person to look for a sheltered place, such as shelters, but also bridges, porches, or recreational centers, resulting in greater opportunities for social interaction. Also, the fraction of rainy days in one's first year on the street/shelter/slum takes into account a potentially higher level of effort in searching for homeless friends at the beginning of one's homeless status. Presumably, homeless people are more likely to look for friends at the beginning of their spell of homelessness, since they need information to survive on the street, such as those regarding soup kitchens and the locations of showers As in Conti et al. (2012), the number of friendship nominations in the first year as a homeless person can be interpreted as a measure of the early accumulation of personal social capital. Hence, the prediction is $\theta > 0$. The variation in $Rain_{ipm}$ comes from the different date (day/month/year) on which a given individual arrived on the street. The date of arrival is specific to the individual and depends on economic and family related shocks during a person's life cycle. It is therefore plausibly as good as random. Those who arrived on the street for the first time in the same day, month and year have been exposed to the same quantity of rainfall.

Before explaining the variable $FormerInmates_{ipm}$, let's analyse potential concerns that could threaten the validity of the instrument. The key identifying assumption is that, conditional on other covariates, the fraction of rainy days in the first year as a homeless person cannot directly influence the probability of imprisonment over the entire spell of homelessness, but neither is it correlated with any other (omitted) factor that could affect $Prison_{ipm}$. More formally, $Cov\left(Rain_{ipm}; \varepsilon_{ipm}|X_{ipm}\right) = 0$.

Different checks have been carried out to test the validity of the exclusion restriction. First, the results presented in Table A1 in the Appendix shows no correlation between $Rain_{ipm}$ and the probability of being in prison before homelessness, a variable potentially correlated with the probability of imprisonment during homelessness (see Panel B, columns 1-2). Second, I look at the relationship between rainfall and crime using two alternative sources of data. In particular, I study the association between the monthly number of individuals arrested, as obtained from the Milan authorities, and the average monthly number of rainy days in Milan from 1998 to 2007: figures indicate a fairly low (.02) and not statistically significant pairwise correlation (p-value equal to .86). Furthermore, I look at the number of crimes reported by the police to the judiciary authority from 1983 to 2003 in

Milan. These figures are published yearly by the Italian National Institute of Statistics (ISTAT), and they allow a disaggregation by province 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 -.32, but, once again, not statistically significant (p-value equal to .16). With the same source of data, I focus on a series of crime categories that are the most common among the homeless: property crimes (burglary, robbery, common theft), drug-related crimes, and offenses related to prostitution: the p-value, equal to .18, confirms a non-statistically significant correlation between the number of rainy days and the number of crimes most commonly committed by homeless people.

Finally, to further rule out concerns about the validity of the exclusion restriction, I look at the evidence coming from previous studies. The relation between weather and crime has been broadly discussed in criminology. This relation appears to greatly vary with the weather (i.e. temperature, rain, wind, and humidity) and the type of crime examined (DeFronzo, 1979; Perry and Simpson, 1987; Cohn, 1990). In particular, previous research that has investigated the relation between rainfall and crime did not find any statistically significant correlations between precipitation and property crime, the homicide rate, rapes, or assaults (Feldman and Jarmon (1979), DeFronzo (1979), Perry and Simpson (1987)). Although these studies did not allow any unambiguous inferences about cause and effect, they seem to suggest that high temperature, more than rainfall variability, might directly influence the individual's decision to commit a crime. For example, violent crimes, such as assaults, burglary, violence, and rape, tend to increase with the ambient temperature at least up to 85°F. A potential explanation is that high temperature may influence changes in social behaviour, such as alcohol consumption, which may be a contributing factor to many violent crimes (Cohn, 1990). In line with these findings, in the economic literature, Jacob, Lefgren and Moretti (2007) examine the short-run dynamics of criminal behaviour by using weather shocks, namely temperature and rainfall, to instrument the level of crimes in a jurisdiction in the US: they show that while temperature is strongly correlated with property crime, the coefficient on precipitation is not statistically significant. Similarly, Ranson (2014) shows that temperature but not rainfall has a strong positive effect on criminal behaviour in the US.

A second concern related to the validity of the exclusion restriction is that rainfall may influence the intensity of non-criminal activities, and therefore indirectly affect crime. For example, if rainfall affects income for those who beg on the street, the criminal behaviour of individual i may also be affected.[22] None of the respondents declare begging as their first source of income and only 0.21%

[22]For example, there could be a negative relation between income and rainfall if during rainy days the homeless

declare begging as their second source. Hence, the results do not change by estimating Equation (1) by excluding individuals who declare begging as their source of income.

Taken together, these pieces of evidence suggest that rainfall is unlikely to be directly correlated with the probability of imprisonment or the error term, and is a valid candidate for instrumenting the size of the social networks of the homeless.

[Insert figure 4]

In Figure 4, Panel A, I plot the fraction of rainy days in one's first year as a homeless person by the respondent's duration on the street/shelter/slum.[23] Recall that (1) includes controls for the month of entry into the spell of homelessness and for a dummy variable indicating a duration shorter than a year.

A second problem in estimating Equation (1) with a standard OLS model concerns the potential sorting of past criminals and potential criminals into 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 exogenous policies to instrument the share of criminal peers as follows:

$$ShareCriminalsFriends_{ipm} = \alpha + \eta FormerInmates_{ipm} + \theta Rain_{ipm} + zX_{ipm} + \rho_p + \varsigma_m + u_{ipm} \quad (3)$$

where $FormerInmates_{ipm}$ is the monthly number of former inmates released from correctional facilities in Milan during i's first year as a homeless person. More precisely, it is computed by dividing the monthly number of former inmates in one's first year as a homeless person by the number of months one has been homeless (i.e. 12 months for respondents with a duration of homelessness longer than one year, but the number of months on the street/shelter/slum for respondents with a duration of homelessness shorter than one year). In Figure 4, Panel B, I show the variation in the share of former inmates by duration of spell of homelessness. The idea behind the instrument is that exogenous policies driving the outflow of inmates increase the supply of criminal potential friends and, consequently, positively affect the likelihood of meeting more criminal peers. The outflow of inmates

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.

[23] The maximum value for the fraction of rainy days (.9) is recorded by homeless people 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 recorded in 6 weather stations 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.

depends on decisions taken by the Italian Ministry of Justice or on the length of each prisoner's criminal sentence, probably agreed on many years before the release.[24] I rely on the assumption that the share of released inmates during i's first year as a homeless person is neither directly correlated with their probability of being imprisoned nor correlated with any other factors that could affect $Prison_{ipm}$. More formally, $Cov(FormerInmates_{ipm}|X_{ipm}; \varepsilon_{ipm}) = 0$. The evidence reported in Panel B in Table A1 in the appendix is in line with this assumption: the coefficient on $Prison^{Before}$ does not affect the probability of imprisonment.

To better understand in which circumstances the instruments play a role, Figure A1 in the Appendix shows all the possible combinations between the time of imprisonment for the individual i, a potential criminal friend j and the time of homelessness between i and j. Indeed, in interpreting β and γ as social interaction effects, we are implicitly assuming that a potential peer influences i's criminal behaviour because their imprisonment happened before that of the respondent and the meeting between the two happened before the respondent's imprisonment.

Let's start by looking at scenarios A and B in Figure A1. In both cases, i went to prison after j and, independently of the fact that i became homeless before or after j (scenario A and B, respectively), we are in a situation of peer exposure, where indeed the realization of the dependent variable in Equation (1) happened after the realization of the two endogenous variables. An alternative hypothesis within scenarios A and B would be if the meeting between i and j happened after i's imprisonment. In this case, the meeting (or friendship) will be endogenous, since, for example, it could have been driven by the individuals' characteristics. By instrumenting the number of friends, we will address the endogeneity issue. In scenario C, i became homeless after j and, in principle, their incarceration should have happened before j's incarceration. But since in Equation 1 we are only considering i's imprisonment during their homelessness (and not before), i's period in prison cannot, by construction, happen before j's period in prison. Scenario C is therefore impossible in our setting. Finally, scenario D could be another case in which the IV strategy plays a role. Indeed, i's imprisonment happened before j's imprisonment, and i and j became friends after being in prison. Therefore their friendship may be driven by self selection and, again, γ would be endogenous. By instrumenting the share of criminal friends, I attempt to address this issue.

[24] Exogenous policies driving the release of prisoners in Italy have been already exploited in the literature. Barbarino and Mastrobuoni (2013) exploit a series of sudden collective pardons enacted in Italy 1962–1990 to estimate the 'incapacitation effect'. Drago, Galbiati and Vertova (2009) and Drago and Galbiati (2012) exploit the 2006 Italian prison pardon to evaluate the role of expected punishmnet and peer effects in criminal behavior.

[Insert figure 5]

Another plausible concern related to the estimation of Equation (1) may arise for the sub-sample of respondents whose homelessness lasted less than a year. For these individuals, the instruments–the fraction of rainy days and the fraction of former inmates—are functions only of the duration of the homelessness. In this subsample, the identification of β and γ comes from the fact that the direct effect of the duration on the outcome variable may be smooth relative to the non-linearity in the instruments. In Figure 5, I plot the duration of the spell of homelessness and the two instruments against the probability of imprisonment. The upper panel does not include the controls, while the lower panel includes the controls as specified in Equation (1). The figures clearly show a linear correlation between imprisonment and the duration of the spell of homelessness, but the correlation between the instruments and the outcome variable shows no apparent pattern.

Finally, an important data-related issue arises because our estimates are based on a stock of individuals who were homeless on a particular date. We do not observe any of those homeless who were in jail at the date of the survey (for example, because they were less exposed to rainfall and so more likely to have a lower number of friends) and this may generate a sample selection problem. In particular, the probability of being in the sample depends on the time spent in that state, and it is therefore higher for individuals with longer windows of homelessness (length-biased sampling) (see Salant, 1977; and Cameron and Trivedi, 2005). High frequency criminal individuals (or individuals with fewer friends) may be imprisoned more often, and, on average, may have shorter spells of homelessness. As a consequence, they may be less likely to be represented in the stock population. Hence, our parameter estimates in Equation (1) are based on a sample of less criminal individuals compared to the true population, with a consequent bias compared to the true effect.

5 Results

5.1 First-stage results

The results presented in this section examine how the number of friends and the friends' characteristics influence the probability of imprisonment for individual i during their period of homelessness. In all estimates, the standard errors are robust and adjusted for clustering of the residuals at the address of interview level (85 clusters).[25] Summary statistics on the variables used in the regressions are described in Table 2.

[25] In table A6 in the appendix I show the main results reported in table 7 by clustering the standard errors at the month by year level.

[Insert table 5]

I begin by analyzing the impact of rainfall on friendship. Table 5 presents results from several first-stage regressions. The outcome variable is the number of nominated best friends, going from 0 to 5. According to the estimates in column 1, the fraction of rainy days during i's first year on the street/shelter/slum strongly and positively predicts the number of friends. During rainy days, there is a higher concentration of homeless people in sheltered places (i.e. bridges, train stations, the underground), with a consequent higher probability of social interactions. In terms of magnitude, an increase by one standard deviation in the fraction of rainy days during one's first year as a homeless person is associated with having 0.50 more friends (column 1). The effect is statistically significant at the 1% level. This finding is robust to the inclusion of a variety of controls, such as age, age squared, gender, education, nationality, length of the homeless spell (in months), a dummy for the length of the spell of homelessness's being less than a year, a dummy indicating the place of the interview (i.e. street or shelter or slum), and dummies for the month of entry into the spell of homelessness (column 3). In column 2, I include the number of former inmates in one's first year as a homeless person. Note that, conditional on the controls, the share of rainy days remains positive and statistically significant at the 1% level, while the variable 'Former Inmates' does not influence the total number of friends. The latter result reinforces the hypothesis that the number of released inmates influences the outcome variable in (1) only through the share of criminal friends and not through the total number of friends. These effects are robust to the inclusion of a second order polynomial for duration (column 4), economic indicators and friends' average characteristics (column 5), a third order polynomial for duration (column 6), and a year of entry dummies. In terms of magnitude, holding the other controls at the sample mean, an increase by one standard deviation in the rainfall increases the dependent variable by approximately 0.5 friends (column 5).

[Insert table 6]

Table 6 presents the first stage results for the share of friends who went to prison already before being homeless on the number of inmates released by Milan's authority during one's first year as a homeless person. The idea behind this instrument is that exogenous policies driving the outflow of inmates affect the supply of criminal potential friends, thus increasing the likelihood of meeting homeless friends who have already had previous criminal experience. According to the baseline estimates in column 1, an increase by one standard deviation in the number of inmates released during one's first year as a homeless person increases the fraction of that homeless person's peers who went to prison already before being homeless by 9 percentage points and the coefficient is statistically significant at the 1 percent level. As in Table 5, we test the robustness of this result by including in Equation (3) the share of rainy days in one's spell of homelessness (column 2), socio-demographic control (column 3), a non-linear effect of duration (columns 4 and 5), economic indicators and friends' characteristics (column 5) and year of entry dummies (column 7). In all the specifications, the number of former inmates is positively and strongly correlated with the share of criminals in one's friendship network. The magnitude of the coefficient increases with the inclusion of the controls: in the specification with economic indicators and friends' characteristics, ceteris paribus, holding the other controls at the sample mean, an increase by one standard deviation in the share of former inmates in one's first year as a homeless person increases the share of one's criminal friends by approximately 11 percentage points (column 5).

In Tables 5 and 6, the F-test and the Cragg-Donald test for weak instruments are well above the critical values of 10 and 7.03 in almost all specifications, suggesting that variation in the share of rainfall and in the number former inmates in one's first year as homeless can be considered valid instruments in this setting (Stock and Yogo, 2002).

Panel A, Table A1 in the Appendix presents reduced form estimates. The coefficients on the share of rainy days (column 1) and on the number of former inmates (column 2) are of the expected signs and statistically significantly correlated with the probability of incarceration, further mitigating concerns about the possibility of weak instruments (Angrist and Kruger, 2001).

5.2 Main results

[Insert table 7]

The main research question of the present paper is to understand the role played by the number of friends and by friends' characteristics in individual criminal decisions. Table 7 presents the second stage estimates where the dependent variable is equal to one if the homeless person i has been imprisoned at least once during their period of homelessness. Columns 1–4 show the OLS results. According to the baseline estimate in column 1, the likelihood of going to prison during homelessness decreases by 3.8 percentage points with one additional friend. A plausible interpretation for this finding is that friends represent a source of insurance. The homeless with more friends have greater chances of surviving on the street without committing crimes because their idiosyncratic shocks are shared among a higher number of individuals.[26] In the next section, I attempt to provide evidence for this possible mechanism. The estimated peer effects, γ, are captured by the share of i's friends who went to prison before becoming homeless. The OLS estimates in column 2 reveal the presence of peer effects on crime: the share of peers who served a previous jail sentence increases the likelihood that an individual, with no prior adjudication, will be imprisoned. The coefficient remains positive and highly significant with the inclusion of socio-demographic controls, place of interview dummies, month of entry dummies, friends' characteristics, and economic indicators (columns 3 and 4). The OLS estimates are consistent with the hypothesis that having more friends reduces the probability of imprisonment, while the share of criminal friends increases it. The estimates in columns 1–4 may, however, simply reflect the effect of unobservables that simultaneously affect the likelihood of having many friends and of being arrested.

Columns 5–10 contain the key results of this empirical section. They show IV estimates of the probability of incarceration. I exploit the fraction of rainy days in one's first year as a homeless person as an instrument for the size of the network, and the variation in the number of former inmates in one's first year as a homeless person as an instrument for the share of one's criminal peers. By looking at the specification with only the total number of friends (column 5), the estimates confirm a negative and statistically significant correlation between the number of friends and the likelihood of imprisonment. Controlling for the endogeneity of γ, in column 6, the estimates clearly indicate the presence of significant peer effects on the probability of committing criminal activities and the coefficient on 'the share of friends in prison' is robust to the inclusion of socio-demographic controls (column 7). Although the evidence in Figure 5 suggests that the effect of duration on the probability of imprisonment is linear, it is very hard to speculate about the functional form of this duration on a priori grounds. I therefore explore the presence of non-linearity in the effect of duration on imprisonment by adding a second and third order polynomial in the duration to Equation (1) (columns 8–9). In all the specifications, the main coefficients of interest maintain the same sign and significance. In particular, in the specification with economic indicators and friends' characteristics (column 9), one additional friend leads to approximately a 12 percentage point decrease in the probability of going to prison during a spell of homelessness, whereas an increase by one standard deviation in the fraction

[26] An alternative interpretation of the negative coefficient on the number of friends could be that having more friends reduces the probability of imprisonment, conditional on crime. This would be the case if, for example, the homeless with larger social networks were to have a lower chance of being caught by the police—because protected by a higher number of friends—compared to those with smaller 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 the credentials to be trusted by the police and have no incentive to be involved in a trial as witnesses to protect a friend.

of peers who served previous jail sentences increases the likelihood of committing illegal acts for an individual with no prior criminal experience by approximately 23 percentage points. These results continue to hold good when including additional socio-demographic characteristics, place of interview and month of entry dummies, controls for friends' characteristics, and quarterly economic indicators in Lombardy. Recall that in the regression we also included the dummy $Prison^{Before}$ to gauge peer effects only on those individuals who did not have prior criminal experience. The inclusion of $Prison^{Before}$ and the length of the spell of homelessness among the controls is important for increasing 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 window of homelessness. Note that by including a dummy for Italians and for the other main nationalities, I attempt to control for potentially different costs of committing illegal acts for Italians and immigrants (i.e. illegal immigrants might face a higher cost of crime because it involves deportation). In the last column of Table 7, I show that the findings are robust to the inclusion of year of entry dummies.[27]

The difference in the magnitude of the coefficients estimated with OLS and IV may be due to the local nature of the IV estimates (Imbens and Angrist (1994)). In this setting, IV estimates can indeed be interpreted as local average treatment effects (LATE) of the number of friends (or the share of criminal friends) on crime for the sub-sample of compliers that are induced to have more friends (or to have more criminal friends) because of a higher share of rainfall (or a higher number of released inmates) in their first year as a homeless person. To characterize these compliers, I estimate first stage Equations (2) and (3) for two mutually exclusive sub-groups: those with a length of spell of homelessness below or above the mean (see Angrist and Fernandez-Val (2013) and Akerman, Gaarder, Mogstad (2015)). The estimates of θ and η are shown in the third column of Table A2 in the Appendix (in Panel A and Panel B, respectively). Following Akerman et al. (2015), I then compute the proportion of compliers in each sub-group as the ratio between θ (and η) in each sub-group and the first stage coefficient in the overall sample, multiplied by the proportion of the sample in the each sub-group (column 2). We see that homeless people with shorter spells are over-represented among the compliers compared to the sample of the homeless at large. These findings suggest a

[27] The effect of the number of friends on criminal behavior is the same using the sub-sample of homeless with at least one friend. Results not displayed but available upon request.

heterogeneity across individuals in the impact of rainfall and number of released inmates. For this reason, we need to be cautious in comparing the OLS and IV estimated coefficients.

[Insert table 8]

The main results presented 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 any peer effects. In Table 8, I use an alternative endogenous variable: a dummy variable equal to one if the respondent reports the name of at least one criminal friend. According to both the OLS and IV estimates, 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 the number of friends. Looking at the magnitude of the coefficient in column 6, an increase by one standard deviation in the one-criminal peer variable leads to a 33 percentage point increase in the likelihood of imprisonment.

Overall, the results in this section show that homeless individuals with many friends are less likely to be imprisoned, but having friends with prior criminal experience increases the likelihood of imprisonment.

5.3 Mechanisms of interactions

To explore the mechanism of social interactions, I now turn to estimating the impact of different 'types' of friends and of the size of the entire friendship network (i.e. friends of friends) on criminal behaviour. More specifically, in this section, I will address the following questions: (i) Do 'old' friends have a stronger influence on one's criminal behaviour? (ii) Are reciprocal relationships more likely to affect the outcome variable? and (iii) Do friends of friends also play a role in shaping an individual's criminal behaviour?

Columns 1–4 in Table A3 in the appendix report OLS coefficients of the effect of having 'old' friends, i.e. nominated friends with spells of homelessness greater than or equal to that of the respondent, and of having 'old' criminal friends, on the probability of imprisonment. Columns 5–10 present the effect of having 'new' friends, i.e. those with spells of homelessness shorter, by at least one year, than the respondent's spell, and of having 'new' criminal friends, on the dependent variable. It is very interesting to note that the coefficient on the number of old friends is negative and statistically significant at the 1% level (column 1) while the coefficient on new friends does not have any effect on the probability of imprisonment (column 5). This finding points in the direction of the previously suggested insurance mechanism: the homeless with long-term friends have a higher probability of being insured against shocks because in long-term relationships, individuals have more chances of reciprocating transactions in the future and of developing a higher degree of trust. behavioural evidence supports the idea individuals better cooperate in infinitely repeated games (Coate and Ravaillon, 1993). Following this argument, a larger network of old friends increases the probability of surviving on the street without committing a crime. Furthermore, columns 2–4 show that the share of old friends who have had previous criminal experience more strongly influences the probability of imprisonment than the share of new criminal friends: in the specification with controls (column 4), an increase by one standard deviation in the fraction of old peers (new peers) who have served a previous jail sentence increases the likelihood of committing an illegal act for an individual with no prior criminal experience by approximately 23 percentage points (11 percentage points).[28]

A great advantage of the Milan Homeless Survey is that one can reconstruct the whole geometric structure of the friendship network so as to (i) investigate which individual characteristics drive the networks' formation, and (ii) more deeply analyse the effect of different types of relationship (e.g. reciprocal versus non-reciprocal) on crime. Figure A2 in the appendix represents a homeless friendship network in a graph which consists of a set of nodes $N = \{1, ..., n\}$ for each homeless person interviewed, and an $n \times n$ matrix g (referred to as an adjacency matrix) describing the friendship network: $g_{ij} = 1$ if i and j are friends, and zero otherwise. Lines between nodes indicate friendship relationships and the arrows describe the direction of the friendship (nominator or nominee). Double arrows represent mutual nominations. To highlight the clustering of criminal behaviour, each node is colored according to the person's imprisonment status: dark blue if the respondent has been in prison at least once, and light blue if the respondent did not have any criminal record.

Thus, we can next test whether the main findings of Table 7 are stronger or weaker for reciprocal nominations. Table A4 in the appendix shows that the number of non-reciprocal friends significantly affects the probability of imprisonment while the coefficient on reciprocal friends does not influence the dependent variable. At first glance, this finding may seem counter-intuitive. However, a plausible interpretation is that adverse income shocks are highly correlated in stronger (reciprocal) relationships while non-reciprocal friends may be more likely to face idiosyncratic shocks. If this is the case, insurance mechanisms among homeless will then better work among non-reciprocal friends (Fafchamps and Gubert, 2007).

[28] Note that by estimating the effect of the number of old and new friends on the probability of imprisonment in the same specification we obtain a very similar result: in a join specification, the coefficient on old friends is equal to -0.034*** and the one on new friends is equal to -0.025.

Finally, Table A5 in the appendix explores the effect of a homeless friendship network on crime. Specifically, I estimate Equation (1) by substituting the number of nominated friends ($NFriends_{ipm}$) with the respondent's network size up to the second degree of separation (best friends and friends of friends) (columns 1–4) and the third degree of separation (best friends, friends of friends, and friends of friends of friends) (columns 5–8). I stopped at the third degree of separation to maintain some individual variation in the network size, given that the majority of the respondents are part of the same network (as shown in Figure A2). At the second degree, the largest network size is equal to 30 friends (mean = 3.3, sd = 4.6) and at the third degree the largest network size includes 59 homeless (mean = 4.7, sd = 7.7). Similarly, the share of criminal friends ($ShareCriminalsFriends_{ipm}$) among best friends has been replaced with the share of criminal friends in the network at the second and third degrees of separation (mean = 0.09, sd = 0.21 and mean = 0.08, sd = 0.20, respectively).

The specifications in columns 1-8 show that the size of the network, both at the second and the third degree of separation, negatively influences the probability of imprisonment. However, the effect of network size on one's probability of imprisonment is statistically significant only until the second degree of separation (column 3-4) and the magnitude of the coefficient on network size decreases as the number of degrees of separation increases. So, more individuals create a better insurance probability (as found in Table 7), but the results show that this effect diminishes as the degree of separations within one's friendship network increases. This finding is in line with many previous studies on social networks (see for example, Fowler and Christakis, 2008). Finally, the estimated coefficients in columns 2–4 and 6–8 seem to show that a higher fraction of peers who served a previous jail sentence increases the likelihood of committing an illegal act.

Tables A3-A5 show interesting results underlying potential mechanisms behind social interactions. However, these findings need to be interpreted with caution, given the endogeneity problem potentially present in OLS estimation.

5.4 A simple simulation exercise

To gain more insight into the potential policy implications of these results, I have computed a simple exercise comparing an increase in the number of friends by one extra criminal friend versus one extra non-criminal friend, on the probability of imprisonment, given the average number of friends and the average fraction of peers in prison before becoming homeless in the sample.

Recall that the number of criminal friends is equal to the share of criminal friends (0.23 as shown in Table 2) multiplied by the total number of friends (1.49 as shown in Table 2), which results in 0.34.

Let's start by increasing the average number of friends by one additional non-criminal friend, from 1.49 friends to 2.49 friends. The share of criminal peers will then decrease from 0.23 to 0.14.[29] The variation in the average share of criminal friends, with and without one extra non-criminal friend, is then equal to -0.09. Based on the coefficients on the number of friends and on the fraction of criminal peers estimated in Table 7 column 9, we can now compute the probability of imprisonment given an extra non-criminal friend. It turns out that, ceteris paribus, the likelihood of imprisonment decreases with an extra non-criminal friend by approximately 19 percentage points.[30]

We can now do the same exercise by calculating the effect of having one additional criminal friend on the probability of going to prison. The total number of criminal friends will be 1.34 and the total number of friends will be 2.49. Consequently, the share of peers in prison before being homeless will increase to 0.54 (1.34/2.49). Looking at the estimation results in Table 7, it turns out that one additional criminal friend increases the likelihood of imprisonment by about 27 percentage points.[31]

Finally, we can compute the expected value of the probability of going to prison by using the variation with one additional criminal and non-criminal friend and, as a proxy for the probability of being a criminal, the average number of offenders in the sample (26%). It turns out that the expected value of going to prison decreases by 8 percentage points with one additional friend, independently of their criminal record before homelessness.[32] This implies that boosting friendship would be an important policy for reducing criminal behaviour among the homeless.

6 Conclusions

This paper uses unique data on the homeless to analyse the role played in their criminal behaviour by the number of friends and by criminal friends. To control for the non-random matching of individuals to their peers, I exploit across-individual variation in the timing of the beginning of each spell of homelessness: I use rainfall shocks in one's first year as a homeless person to instrument the size of the network, and variation in the number of inmates released per month by the Italian authorities to proxy the fraction of criminal friends in the network. The results provide strong evidence for peer effects in the realm of criminality among homeless individuals. Specifically, the probability of being arrested during a spell of homelessness increases with the exposure to peers with criminal records.

[29] 0.13 is computed by dividing the number of criminal friends by the number of friends. That is 0.34/2.49=0.14.

[30] This result is obtained by substituting 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 (1) as follows: $Prison_i$ =1 (-0.120)+(-0.09) (0.726)=-0.19.

[31] The likelihood of imprisonment has been computed as follows: $Prison_i$=1(-0.120)+(0.54) (0.726)=0.27.

[32] The likelihood of imprisonment has been computed as follows: $Prison_i = (0.26)(0.27)+(1-0.26)(-0.19) = -0.08$.

However, the same probability decreases as the total number of homeless friends (criminal and non-criminal) increases. A simple simulation exercise shows that among these two opposite results, the latter effect dominates.

These results may have some implications for policies. First, the existence of peer effects on criminal behaviour suggests that any reduction in the criminal activities of one's friends would lead to further reductions in crime. Hence, programs aiming at rehabilitation in prison and housing assistance after release may have a beneficial spillover in reducing former inmates at risk of becoming homeless and, consequently, in decreasing the propagation of crime among the homeless. Furthermore, the simulation exercise shows that the probability of imprisonment during a spell of homelessness decreases with one additional friend, independently of their criminal history. This finding may provide an incentive to policy makers to boost social interactions among the homeless. For instance, drop-in day centers that offer social activities for homeless people, such as board games, group conversations, and movies, may be beneficial for stimulating social interactions and reducing criminal activity.

Could these results and policies be extended to other settings or populations? Given that homelessness is a critical consequence of the economic crises, researchers should carefully reflect on the thin line between the urban poor and the homeless. Certainly, the homeless are one of the most vulnerable populations in the richer countries, with a high degree of social interactions and a high share of former inmates.

References

[1] Akerman A., Gaarder I., Mogstad M. (2015). 'The Skill Complementarity of Broadband Internet', NBER Working Paper, 2015

[2] Anderson L., Snow D. A. (1993). 'Down on Their Luck. A Study of Homeless Street People', University of California Press.

[3] 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.

[4] Angrist J., Fernandez-Val I. (2013). 'ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework', *Advances in Economics and Econometrics*, Vol. 3.

[5] Allgood S., Warren R. S., (2003). 'The duration of homelessness from a national survey', *Journal of Housing Economics* 12, 273–290.

[6] Barbarino A. and Mastrobuoni G. (2013). 'The Incapacitation Effect of Incarceration: Evidence from Several Italian Collective Pardons', *American Economic Journal: Economic Policy.*

[7] Bayer, P., Hjalmarsson R., Pozen D., (2009). 'Building criminal capital behind bars: Peer effects in juvenile corrections', *Quarterly Journal of Economics*, 124, 105–147.

[8] Becker, G., (1968). 'Crime and Punishment: An Economic Approach'. *The Journal of Political Economy* 76, 169–217.

[9] Bentolila, S. and Ichino, A. (2007). 'Unemployment and Consumption. Near and far away from the Mediterranean?', *Journal of Population Economics,* 21, 225–289.

[10] Black S, Devereux P, Salvanes K. G. (2013). 'Under pressure? The effect of peers on outcomes of young adults', *Journal of Labor Economics*, 31 (1), 119–153.

[11] Bobonis G., Finan F. (2009). 'Neighborhood peer effects in secondary school enrollment decisions', *Review of Economics and Statistics*, 91 (4), 675–716.

[12] Braga M., Corno L. (2012). 'Being homeless: Evidence from Italy', *Giornale degli Economisti e Annali di Economia,* 21 (3), 465–498.

[13] Brent B. (2007). 'A Repeated Observation Approach for Estimating the Street Homeless Population', *Evaluation Review*, 31, 166.

[14] Burt M. (1992). 'Over the Edge: The Growth of Homelessness in the 1980s', *Urban Institute Press*, Washington, DC.

[15] Calvò-Armengol A., Patacchini E., Zenou Y. (2009). 'Peer Effects and Social Networks in Education', *Review of Economic Studies*, 76, 1239–1267

[16] Calvò-Armengol A., Zenou Y. (2004). 'Social network and crime decisions: The role of social structure in facilitating delinquent behaviour', *International Economic Review,* 45, 939–958.

[17] Cameron A. C., Trivedi P. K. (2005). 'Microeconometrics. Methods and Applications', Cambridge University Press.

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

[19] Coate S. and Ravaillon M. (1993). 'Reciprocity without committment: characterization and performance of informal insurance arrangments', *Journal of Development Economics,* 40:1-24.

[20] Cohn, E. G. (1990). 'Weather and Crime', *British Journal of Criminology*, 30 (1), 51–64.

[21] Commissione d'indagine sull'esclusione sociale, 2002. *Rapporto sulle politiche contro la povertà e l'esclusione sociale 1997–2001*, Carocci, Roma.

[22] Conti G., Galeotti A., Mueller G., Pudney S. (2012). 'Popularity', *Journal of Human Resources,* 48 (4), 1072–1094.

[23] Cook P., Gross K. A. (1996). 'A selective review of the social contagion literature', Terry Stanford Institute Working Paper, Duke University.

[24] Damm A. P., Dustmann C. (2014). 'The Effect of Growing Up in a High Crime Neighborhood on Criminal behaviour', *American Economic Review*, 104 (6), 1806–1832.

[25] De Fronzo J. (1984). 'Climate and crime: Tests of an FBI assumption', *Environment and behaviour*, 16, 185–210.

[26] De Giorgi G., Pellizzari M., Redaelli S. (2010). 'Identification of Social Interactions through Partially Overlapping Peer Groups', *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*, 47 (8), 1152–1177.

[28] Dercon S., De Weerdt J. (2006). 'Risk sharing networks and insurance against illness', *Journal of Development Economics* 81, 337–356.

[29] Drago F., Galbiati R., Vertova P. (2009). 'The Deterrent Effects of Prison: Evidence from a Natural Experiment, *Journal of Political Economy*, 117 (2), 254–280.

[30] Drago F., Galbiati R. (2012). 'Indirect Effects of a Policy Altering Criminal behaviour: Evidence from a Natural Experiment', *American Economic Journal: Applied Economics,* 4 (2), 199–218.

[31] European Commission. (2010). *A Strategy for Smart, Sustainable and Inclusive Growth, Communication from the Commission, Europe 2020,* March 3, 2010.

[32] Fafchamps M., Gubert F. (2006). 'The Formation of risk sharing networks', *Journal of Development Economics,* 83 (2), 326–50.

[33] Fafchamps M., Gubert F. (2007). 'Risk Sharing and Network Formation', *American Economic Review Papers and Proceedings*, 97(2): 75-79

[34] Fafchamps M., Lund S. (2003). 'Risk Sharing Networks in Rural Philippines', *Journal of Development Economics,* 71, 261–87.

[35] Fafchamps M., Blumenstock J., Eagle N. (2015). 'Risk sharing and mobile phones: Evidence in the aftermath of natural disaster', working paper.

[36] Feldman H., Jarmon, R. (1979). 'Factors Influencing Criminal behaviour in Newark: A Local Study in Forensic Psychiatry', *Journal of Forensic Science,* 24: 234–239.

[37] Fisher N., Turner S., 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.

[38] Fowler J. H., Christakis N. A. (2008). 'The Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis Over 20 Years in the Framingham Heart Study' *British Medical Journal,* 337, a2338; doi:10.1136/bmj.a2338

[39] Glaeser E., Sacerdote B., Scheinkman J. (1996). 'Crime and Social Interactions', *Quarterly Journal of Economics*, 111 (2).

[40] Godlonton S., Thornton R. Forthcoming. 'Peer effects in learning HIV results', *Journal of Development Economics.*

[41] Honig M., Filer R. K. (1993). 'Causes of intercity variation in homelessness', *American Economic Review*, 83 (1), 248–255.

[42] Imbens, G. W., Angrist, J. D. (1994). 'Identification and estimation of local average treatment effects'. *Econometrica*, 62 (2), 467–75.

[43] Italian National Institute for Statistics. (2014). Data available at www.data.istat.it.

[44] Jencks, C., 1994. 'The Homeless'. *Harvard University Press.*

[45] Jacob B., Lefgren L., Moretti E. (2007). 'The dynamics of criminal behaviour: Evidence from weather shock', *Journal of Human Resources,* 42 (3).

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

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

[48] Ligon E., Thomas, J. P. and Worrall T. (2001). 'Informal insurance arrangments in village economies', *Review of Economic Studies,* 69 (1), 209–44.

[49] Liu X., Patacchini E., Zenou Y., Lee L. (2013). 'Criminal networks: Who is the key player?', CEPR Discussion Paper No. 8185.

[50] Lochner L., Moretti E. (2004). 'The effect of education on crime: Evidence from prison inmates, arrest and self-reports', *American Economic Review* 94 (1).

[51] 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.

[52] Manski, C. (1993). 'Identification of Endogenous Social Effects: The reflection problem', *Review of Economic Studies*, 60, 531–542.

[53] Mansur E., Quigley J., Raphael S., Smolensky E. (2002). 'Examining Policies to Reduce Homelessness Using a General Equilibrium Model of the Housing Market', *Journal of Urban Economics*, 52, 316–340.

[54] Moffitt, R. (2001). 'Policy Interventions, Low-Level Equilibria, and Social Interactions', in *Social Dynamics*, Steven Durlauf and Peyton Young, Editors, Cambridge: MIT Press.

[55] O'Flaherty, B. (1996). *Making Room: The Economics of Homelessness*, Harvard University Press, Cambridge, Massachusetts.

[56] Patacchini E., Zenou Y. (2012). 'Juvenile Delinquency and Conformism', *Journal of Law, Economics & Organization*, 28, 1–31.

[57] Perry J. D., Simpson M. E. (1978). 'Violent crimes in a city: Enviromental determinants' *Enviroment and behaviour*, 19, 77–90.

[58] 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.

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

[60] Ranson M. (2014). 'Crime, weather and climate change', *Journal of Environmental Economics and Management,* 67 (3), 274–302.

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

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

[63] Salant S. W. (1997). 'Search theory and duration data: A theory of sorts', *Quarterly Journal of Economics*, 91, 39–57.

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

[65] Sutherland E. H. (1947). 'Principles of Criminology', Fourth Edition, Chicago: Lippincott.

[66] The U.S. Department of Housing and Urban Development, Office of Community Planning and Development. 2008. *A Guide to Counting Unsheltered Homeless People*, Washington, DC.

[67] The U.S. Department of Housing and Urban Development, Office of Community Planning and Development. 2013. *The 2013 Annual Homeless Assessment Report to the Congress.*

Tables

Table 1: Summary statistics on the homeless count

	Street	*Shelter*	*Slums*
Panel A: Number of homeless			
Counted	408	1152	2300
Sampled	408	500	525
Panel B: Percentage of homeless			
Interviewed	34.60	84.00	66.50
Who refused to answer	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: Milan Homeless Survey, January 2008

Table 2: Summary statistics on street, sheltered homeless and slums dwellers

	Obs.	*Mean*	*SD*	*Min*	*Max*
Panel A: Demographics characteristics					
Female	883	0.28	0.45	0	1
Age	879	39.3	14.6	9	83
Italian	883	0.31	0.46	0	1
Romanian	883	0.29	0.45	0	1
Moroccan/Algerian/Tunisian	883	0.1	0.3	0	1
Other Nationalities	883	0.29	0.45	0	1
No education	879	0.08	0.26	0	1
Primary education	879	0.61	0.49	0	1
Secondary education	879	0.30	0.46	0	1
Duration of homelessness (months)	883	83.8	120.3	0.03	1006
Duration of homelessness < one year (months)	883	0.32	0.46	0	1
Single	879	0.39	0.49	0	1
Divorced	879	0.19	0.39	0	1
Married	879	0.34	0.47	0	1
Widow/er	879	0.04	0.19	0	1
Panel B: Imprisonment					
Prison	883	0.26	0.44	0	1
Prison during homelessness	883	0.21	0.40	0	1
Prison before homelessness	883	0.10	0.30	0	1
Prison only during homelessness	883	0.18	0.38	0	1
Prison only before homelessness	883	0.10	0.30	0	1
Prison before and during homelessness	883	0.03	0.16	0	1
Panel C: Friendship					
N. of best friends	883	1.49	1.63	0	5
At least one friend	883	0.63	0.48	0	1
Network size at 2° degree	883	3.25	4.56	0	30
Network size at 3° degree	883	4.73	7.65	0	59
Share of friends in prison before homelessness [a)]	523	0.23	0.31	0	1
At least one friend in prison before homelessness	523	0.27	0.48	0	1
Panel D: Instruments					
Share of rainy days during one’s first year of homelessness [b)]	883	0.23	0.10	0	0.9
Former inmates during one's first year of homelessness [c)]	761	6962.962	1065.82	0	11362

Notes: Sample of respondents living in the street, emergency shelters or in slums. In Panel A, “*Duration of homelessness*” is the number of months from the first day on the street/shelter/slum until the date of the survey. In Panel C, “*the share of friends in prison before homelessness*” has been computed by dividing the number of friends who have been in prison before becoming homeless and the number of friends interviewed in the survey. It is therefore missing for respondents with no friends or who cited the name of friends not interviewed. In Panel D, the "*share of rainy days during one's first year as an homeless person*" is computed as the ratio between the sum of the rainy days in one’s first year as homeless and 365 days for respondents with homelessness duration longer than one year or the number of days on the street/shelter/slum for respondents with homelessness duration shorter than one year; "*Former inmates during one's first year as an homeless person*" is computed by dividing the monthly number of former inmates in one's first year as homeless and 12 months for respondents with homelessness duration greater than one year and the number of months on the street/shelter/slum for respondents with homelessness duration lower than one year. Source: Milan Homeless Survey 2008 for demographic characteristics, friendship and imprisonment; 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 former inmates.

Table 3: Distribution of friends and criminal friends, by place of interview (%)

	All Sample	*Street*	*Shelter*	*Slums*
	(1)	(2)	(3)	(4)
Panel A: Distribution of friends				
0 friends	37.49	31.25	29.56	49.00
At least 1 friend	62.51	68.75	70.44	51.00
1 friend	24.46	29.69	29.8	16.33
2 friends	15.4	17.19	18.47	11.17
3 friends	7.13	6.25	8.37	6.02
4 friends	5.66	7.81	5.67	4.87
5 friends	9.85	7.81	8.13	12.61
Mean	1.49	1.53	1.55	1.39
Observations	883	128	406	349
Panel B: Share of friends in prison before homelessness[a)]				
0	72.08	75.86	76.28	62.96
At least 1 friend in prison before homelessness	28.27	24.14	24.46	37.04
0.20	0.76	--	0.36	2.47
0.25	--	--	--	--
0.33	2.29	3.45	0.36	4.94
0.40	0.38	--	--	1.23
0.50	5.74	3.45	5.11	8.02
0.67	0.76	--	1.09	0.62
0.75	0.38	--	--	1.23
0.80	0.19	--	--	0.62
1.00	17.4	17.24	17.15	17.9
Mean	0.23	0.20	0.21	0.27
Observations	523	87	274	162

Notes: Sample of respondents living in the street, emergency shelters or in slums in January 2008. a) See table 2 for the definition of the share of friends in prison before homelessness. Source: Milan Homeless Survey 2008.

Table 4: Friendship characteristics

	Old Friends	*New Friends*	*Mutual Friends*	*Non-mutual Friends*	
				Nominators	*Nominees*
	(1)	(2)	(3)	(4)	(5)
Panel A: All Sample					
0	55.04	77.01	77.12	52.86	63.44
At least 1	44.96	22.99	22.88	47.14	36.56
1	26.39	15.29	16.08	30.1	23.05
2	10.99	4.87	3.51	9.1	6.61
3	3.85	2.15	1.59	4.99	3.67
4	2.27	0.68	1.13	2.06	1.62
5	1.47	--	0.57	0.88	0.15
More than 5	--	--	--	--	1.47
Mean	0.76	0.34	0.35	0.76	0.65
Observations	883	883	883	681	681
Panel B: Sample of respondents with at least one friend					
0	28.08	63.22	63.41	8.29	55.43
At least 1	71.92	36.78	36.59	91.71	44.57
1	42.21	24.46	25.72	58.57	28.86
2	17.57	7.79	5.62	17.71	7.43
3	6.16	3.44	2.54	9.71	4.86
4	3.62	1.09	1.81	4	1.72
5	2.36	--	0.91	1.71	--
Mean	1.22	0.55	0.56	1.47	0.77
More than 5	--	--	--	--	1.7
Observations	552	552	552	350	350

Notes: Sample of respondents living in the street, emergency shelters or in slums in January 2008. Column 1 describes the distribution of "old" friends and column 2 shows the distribution of "new" friends. "Old" friends are defined as friends with homelessness spells equal or longer than respondent's spell; "New" friends are defined as friends with homelessness spells shorter, at least by one year, than respondent's spell. Column 3 describes the fraction of reciprocal relationships and columns 4 and 5 report nominators and nominees among the sample of non-reciprocal relationships. Source: Milan Homeless Survey 2008.

Table 5: First stage regression - the effect of rainfall on friendship

Dependent variable:	*Number of best friends*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of rainy days	4.971***	4.694***	4.381***	4.337***	4.889***	4.843***	4.437***
	(0.526)	(0.534)	(0.533)	(0.531)	(0.736)	(0.744)	(0.989)
Former inmates [a)]		0.006	-0.030	-0.021	-0.120	-0.130	-0.093
		(0.067)	(0.057)	(0.059)	(0.011)	(0.130)	(0.158)
$\text{Prison}^{\text{Before}}$			0.117	0.113	0.110	0.109	0.131
			(0.201)	(0.202)	(0.208)	(0.209)	(0.194)
Age			0.008	0.005	0.000	0.000	-0.003
			(0.019)	(0.019)	(0.020)	(0.020)	(0.020)
Age squared			-0.000	-0.000	-0.000	-0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female			-0.222	-0.217	-0.197	-0.197	-0.253*
			(0.147)	(0.152)	(0.170)	(0.171)	(0.149)
Primary education			-0.301	-0.291	-0.340**	-0.340**	-0.286*
			(0.184)	(0.184)	(0.166)	(0.165)	(0.170)
Secondary education			-0.461**	-0.450**	-0.499***	-0.493***	-0.449**
			(0.176)	(0.177)	(0.162)	(0.162)	(0.173)
Italian			0.080	0.077	0.030	0.029	0.024
			(0.204)	(0.203)	(0.201)	(0.200)	(0.197)
Moroccan			-0.024	-0.005	-0.009	-0.018	-0.003
			(0.142)	(0.142)	(0.148)	(0.140)	(0.144)
Romanian			-0.092	-0.098	-0.122	-0.122	-0.133
			(0.156)	(0.158)	(0.164)	(0.163)	(0.163)
Duration < 1 year			0.014	0.171	0.105	-0.016	-0.095
			(0.255)	(0.267)	(0.256)	(0.414)	(0.509)
Duration			-0.000	0.008	-0.023	-0.029	0.187
			(0.003)	(0.005)	(0.021)	(0.035)	(0.330)
Duration squared[a)]				-0.041	0.017	0.110	-0.000
				(0.035)	(0.049)	(0.423)	(0.257)
Duration cubic[a)]						-0.000	
						(0.001)	
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes
Month of entry dummies	yes	yes	yes	yes	yes	yes	yes
Friends characteristics	no	no	no	no	yes	yes	yes
Economic indicators	no	no	no	no	yes	yes	yes
Year of entry dummies	no	no	no	no	no	no	yes
Observations	883	483	478	478	477	477	477
F-test	89.13	43.19	35.16	34.35	24.69	22.51	4.36
Cragg-Donald test	--	19.29	23.76	22.25	17.57	13.15	3.95

Notes: Linear regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the address of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in the footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Place of interview dummies are indicators for interviews on the street, in shelters and in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014).

Table 6: First stage regression - the effect of former inmates on criminal friends

Dependent variable:	*Share of friends in prison before homelessness*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Former inmates a)	0.085***	0.100***	0.109***	0.104***	0.106***	0.105***	0.106**
	(0.015)	(0.020)	(0.020)	(0.020)	(0.031)	(0.035)	(0.041)
Share of rainy days		-0.259	-0.377	-0.357	-0.436*	-0.433*	-0.185
		(0.209)	(0.235)	(0.235)	(0.227)	(0.233)	(0.433)
$Prison^{Before}$			0.102	0.104	0.111	0.111	0.080
			(0.067)	(0.066)	(0.069)	(0.069)	(0.073)
Age			-0.003	-0.002	-0.001	-0.001	-0.001
			(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Age squared			0.000	0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female			-0.009	-0.011	-0.007	-0.007	-0.010
			(0.039)	(0.039)	(0.039)	(0.039)	(0.040)
Primary education			0.008	0.004	0.032	0.032	0.046
			(0.068)	(0.068)	(0.068)	(0.069)	(0.064)
Secondary education			0.030	0.025	0.055	0.054	0.077
			(0.071)	(0.071)	(0.071)	(0.072)	(0.071)
Italian			-0.031	-0.030	0.016	0.016	0.058
			(0.063)	(0.063)	(0.072)	(0.072)	(0.076)
Moroccan			-0.019	-0.028	-0.012	-0.011	-0.029
			(0.042)	(0.040)	(0.042)	(0.040)	(0.042)
Romanian			-0.020	-0.018	0.022	0.022	0.044
			(0.073)	(0.073)	(0.077)	(0.077)	(0.076)
Duration < 1 year			0.104*	0.033	0.045	0.054	-0.193
			(0.062)	(0.086)	(0.097)	(0.132)	(0.155)
Duration			-0.001	-0.004*	-0.000	0.000	-0.032
			(0.001)	(0.002)	(0.005)	(0.009)	(0.112)
Duration squared a)				0.010*	0.006	-0.008	0.017
				(0.001)	(0.017)	(0.104)	(0.064)
Duration cubic a)						0.000	
						(0.000)	
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes
Month of entry dummies	yes	yes	yes	yes	yes	yes	yes
Friends characteristics	no	no	no	no	yes	yes	yes
Economic indicators	no	no	no	no	yes	yes	yes
Year of entry dummies	no	no	no	no	no	no	yes
Observations	485	483	478	478	477	477	477
F-test	29.34	19.43	17.70	14.32	6.75	5.89	3.38
Cragg-Donald test	--	19.29	23.76	22.25	17.57	13.15	3.95

Notes: Linear regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the address of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in the footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Place of interview dummies are indicators for interviews on the street, in shelters and in slums. Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014).

Table 7: Second stage - the probability of imprisonment

Dependent variable:	*1 if prison during homelessness*										
	OLS				IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Number of best friends	-0.038***	-0.052***	-0.049***	-0.048***	-0.111***	-0.114***	-0.130***	-0.129***	-0.120***	-0.104***	-0.207**
	(0.009)	(0.009)	(0.008)	(0.008)	(0.029)	(0.026)	(0.034)	(0.033)	(0.039)	(0.039)	(0.091)
Share of friends in prison before homelessness		0.578***	0.557***	0.555***		0.579***	0.582***	0.559***	0.726***	0.838***	0.863*
		(0.055)	(0.057)	(0.056)		(0.181)	(0.149)	(0.151)	(0.206)	(0.206)	(0.516)
Duration < 1 year			0.047	0.032			0.055	0.018	0.064	0.138	0.124
			(0.064)	(0.076)			(0.045)	(0.063)	(0.074)	(0.112)	(0.168)
Duration			-0.003	-0.004			0.000	-0.001	0.001	0.007	-0.124
			(0.003)	(0.004)			(0.000)	(0.002)	(0.004)	(0.006)	(0.084)
Duration squared [a)]				0.005				0.008	-0.002	-0.100	-0.018
				(0.010)				(0.014)	(0.011)	(0.084)	(0.064)
Duration cubic [a)]										0.000	
										(0.000)	
Controls	no	no	yes	yes	no	no	yes	yes	yes	yes	yes
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Month of entry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Friends characteristics	no	no	yes	yes	no	no	no	no	yes	yes	yes
Economic indicators	no	no	yes	yes	no	no	no	no	yes	yes	yes
Observations	883	483	477	477	883	483	478	478	477	477	477

Notes: Linear regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the address of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in the footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Controls include a dummy equal to one if respondent has been in prison before homelessness, age, age squared, female, 2 dummies for educational level (primary education and secondary education and no education is the omitted category), three dummies equal to 1 if the respondent is Italian, Moroccan or Romanian; Place of interview dummies are indicators for interviews on the street, in shelters and in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014). In columns (5) - (10), the number of best friends is instrumented with the share of rainy days during one's first year as an homeless person and the share of friends in prison before homelessness is instrumented with the fraction of former inmates during one' first year as an homeless person.

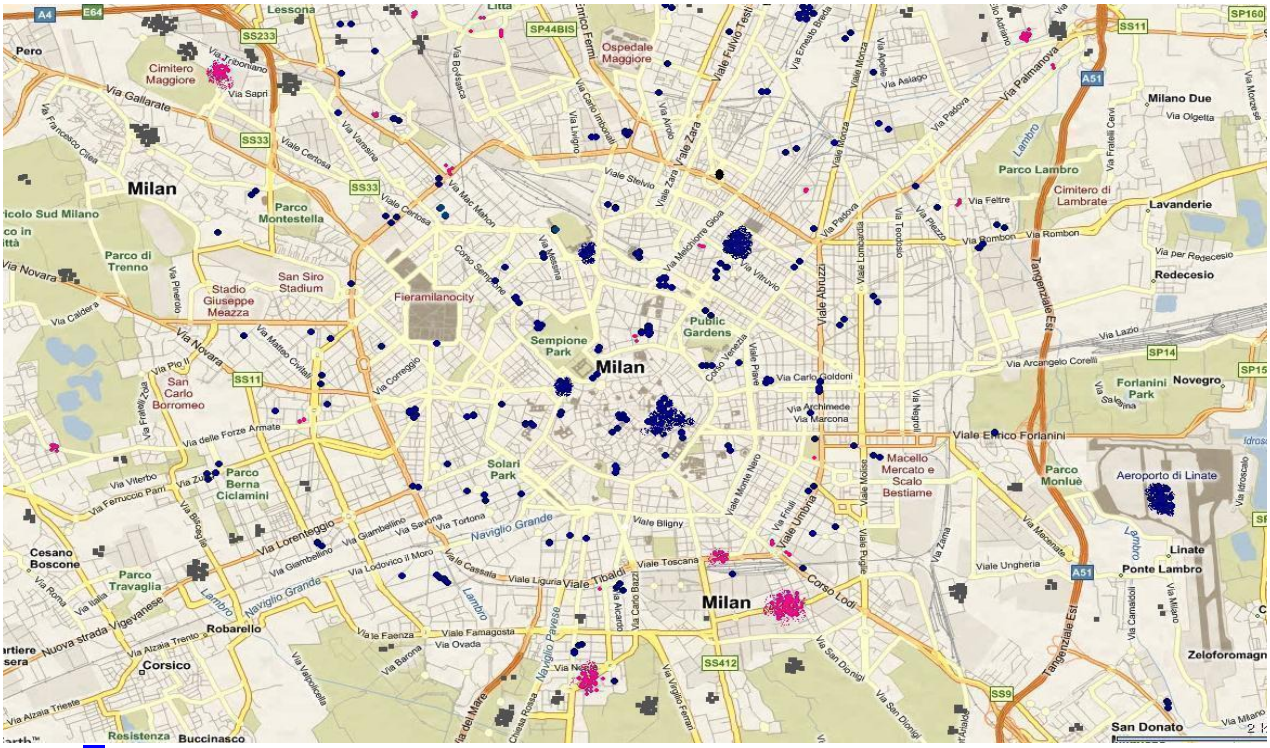
Table 8: Second stage – alternative measure of criminal friends

Dependent variable:	*1 if prison during homelessness*					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of best friends	-0.064***	-0.059***	-0.059***	-0.118***	-0.130***	-0.123***
	(0.009)	(0.009)	(0.009)	(0.025)	(0.033)	(0.042)
At least one friend in prison before homelessness	0.481***	0.467***	0.467***	0.575***	0.545***	0.681***
	(0.053)	(0.049)	(0.049)	(0.163)	(0.141)	(0.215)
Controls	no	yes	yes	no	yes	yes
Place of interview dummies	yes	yes	yes	yes	yes	yes
Month of entry dummies	yes	yes	yes	yes	yes	yes
Friends characteristics	no	no	yes	no	no	yes
Economic indicators	no	no	yes	no	no	yes
Observations	483	478	477	483	478	477
First Stage F-test	--	--	--	44.89	33.43	23.49
First Stage F-test 2	--	--	--	15.86	13.75	7.97
Cragg-Donald test	--	--	--	19.29	16.76	14.43

Notes: Linear regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the address of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in the footnote of table 2. Controls include a dummy equal to one if respondent has been in prison before homelessness, age, age squared, female, 2 dummies for educational level (primary education and secondary education and no education is the omitted category), three dummies equal to 1 if the respondent is Italian, Moroccan or Romanian; a dummy indicating a homelessness spell less than a year, length of the homeless spell, length squared of the homeless spell. Place of interview dummies are indicators for interviews on the street, in shelters and in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014). In columns (4) - (6), the number of best friends is instrumented with the share of rainy days during one's first year as an homeless person and the share of friends in prison before homelessness is instrumented with the fraction of former inmates during one' first year as an homeless person.

Figures

Figure 1: Spatial distribution of street homeless and of shelters and slums' location in Milan



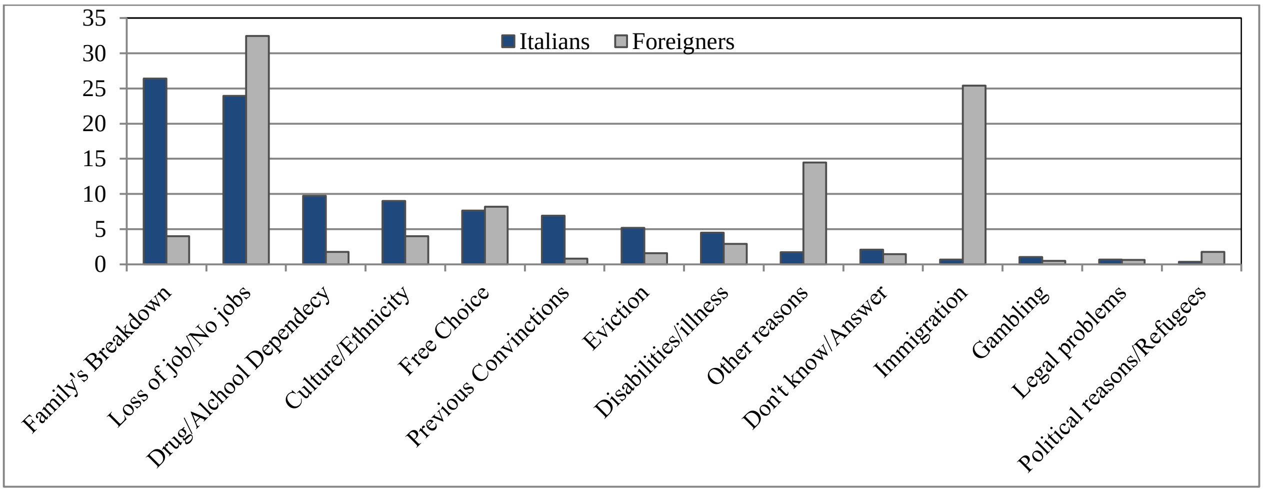
Legend: [ ]=Location of unsheltered homeless, each dot=1 homeless

[ ]= Location of shelters, each dot =10 homeless

[ ]= Location of slums, each dot =10 homeless

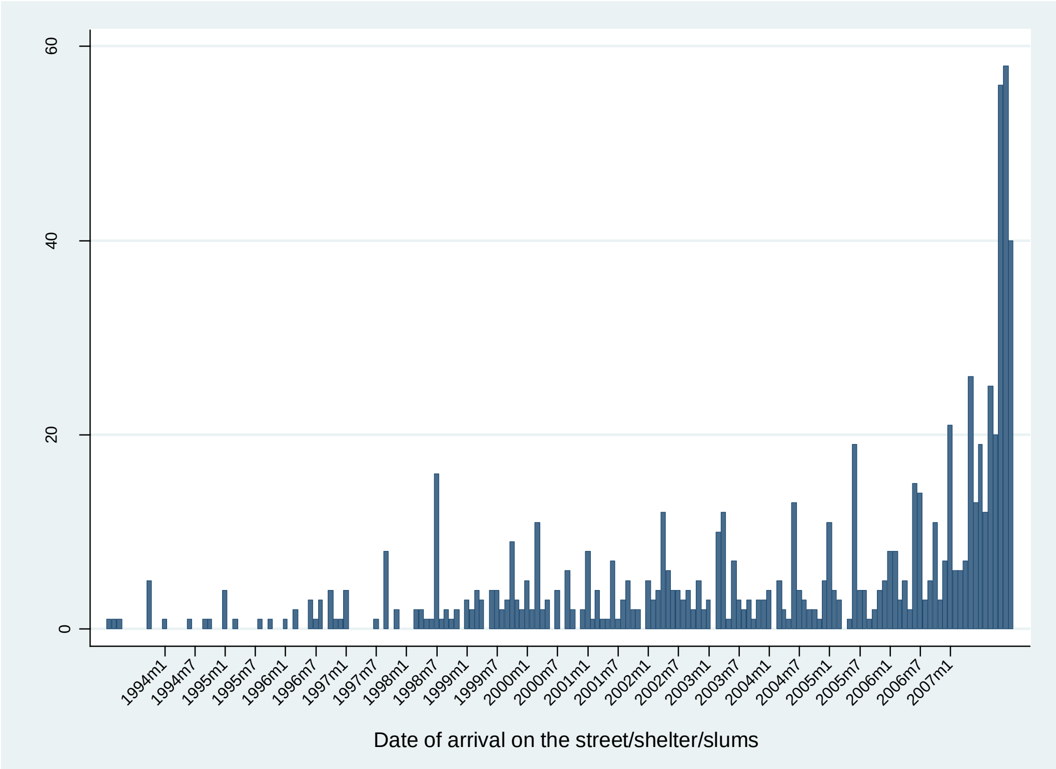
Source: Milan Homeless Survey 2008

Figure 2: Main reason for homelessness, by nationality (%)



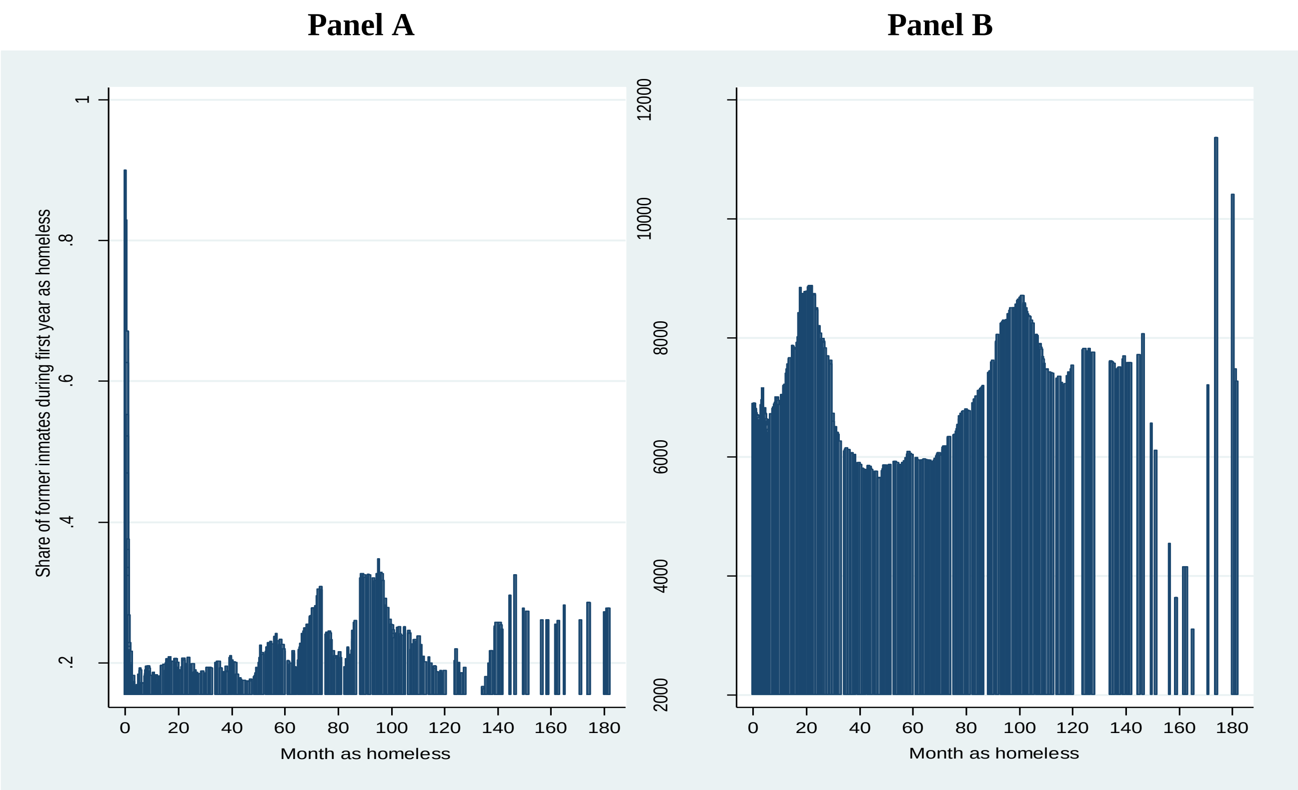
Notes: Sample of respondents living as street, sheltered homeless or in slums in January 2008. Source: Milan Homeless Survey 2008.

Figure 3: Number of homeless, by date of arrival on the street/shelter



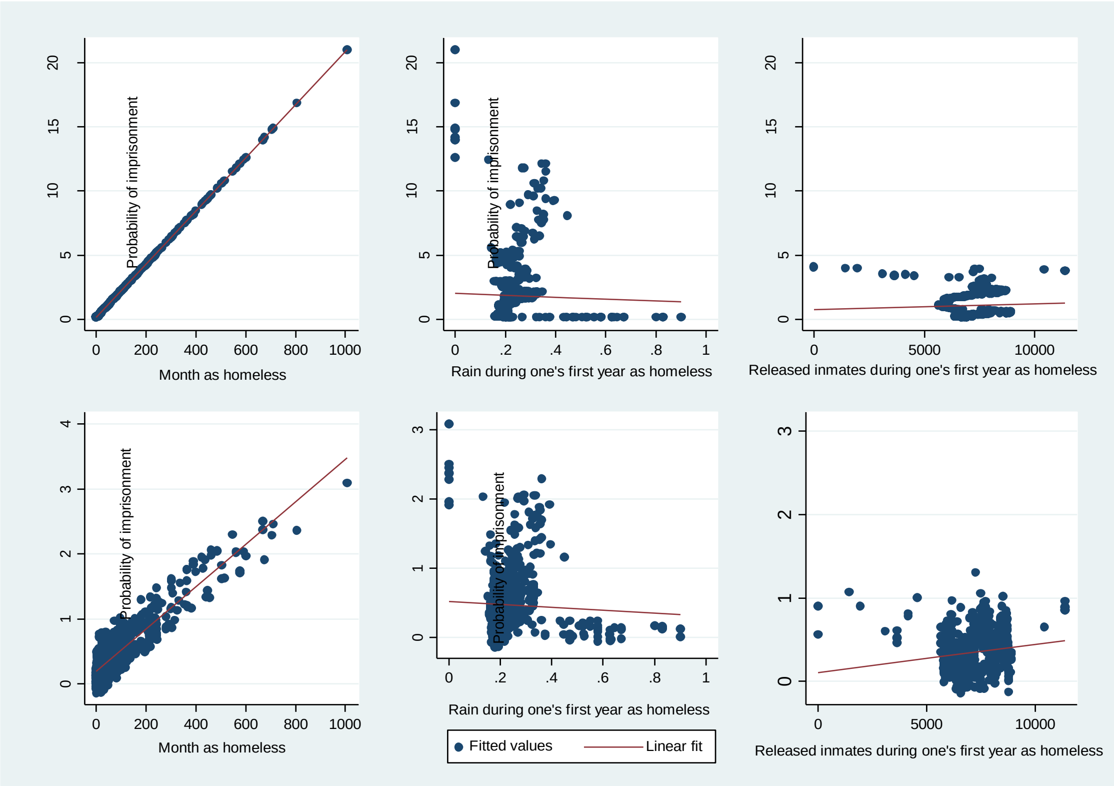
Notes: Sample of respondents living as street, sheltered homeless or in slums in January 2008. Source: Milan Homeless Survey 2008.

Figure 4: Share of rainy days and of former-inmates during one's first year as an homeless person



Notes: Sample of respondents living as street, sheltered homeless or in slums in January 2008. Source: Regional Agency for the Environmental Protection and the Meteorological Department of the Military Aeronautics for the rainfall data, Statistical Offices of the correctional facilities in Milan for data on former inmates.

Figure 5: Relationship between duration, rainfall, former inmates and the probability of imprisonment



Notes: The figures plot the duration of homeless spells, the fraction of rainy days during one's first year as homeless and the fraction of released inmates during one's first year as homeless on the probability of imprisonment for respondents with a homeless spell lower than one year. The first three graphs in the upper panel do not include controls while the three graphs in the lower panel include controls as specified in equation 1. Sample of respondents living as street, sheltered homeless or in slums in January 2008. Source: Regional Agency for the Environmental Protection and the Meteorological Department of the Military Aeronautics for the rainfall data, Statistical Offices of the correctional facilities in Milan for data on former inmates.

On line Appendix

Figure A1: Friends' exposure

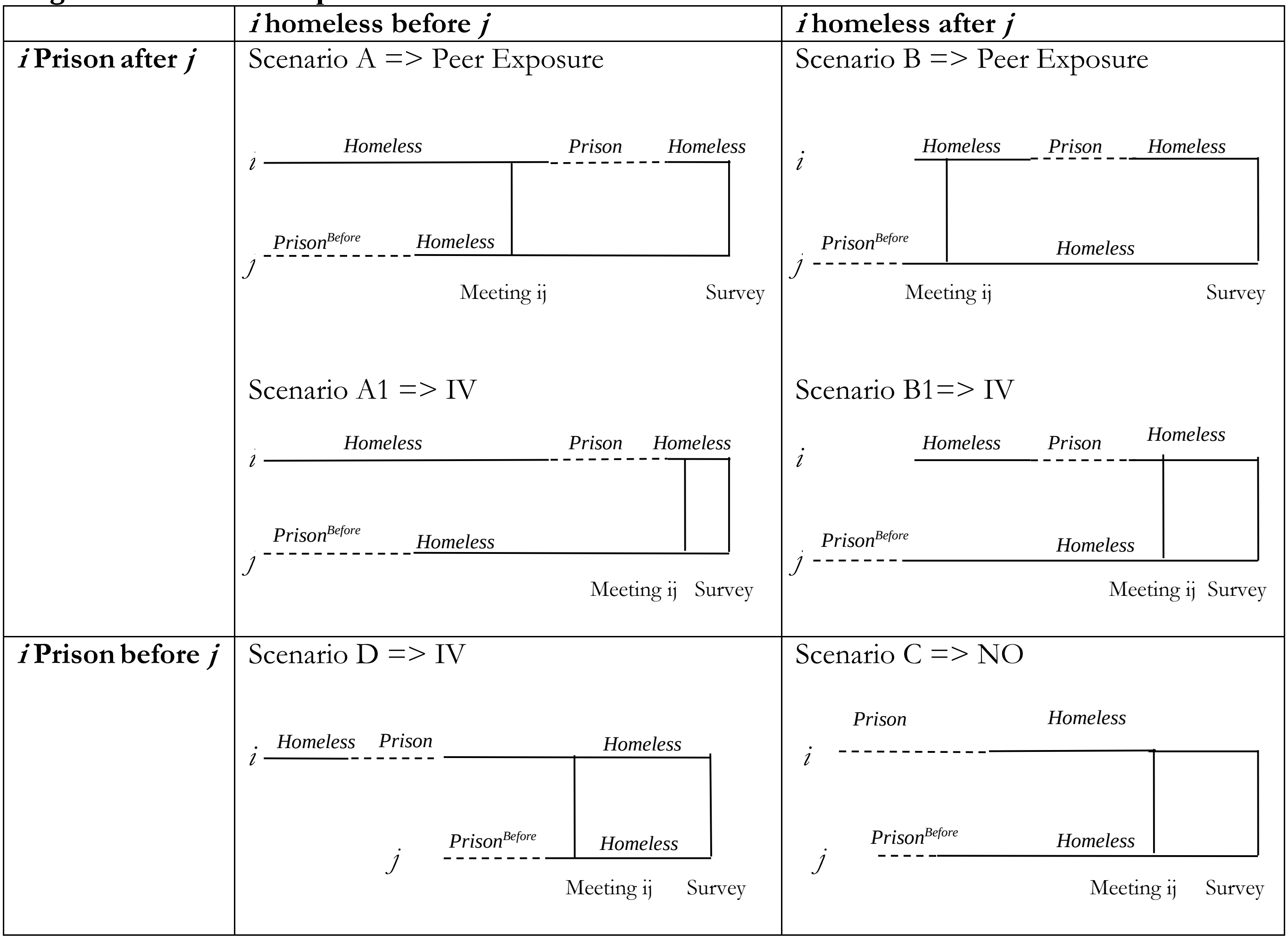
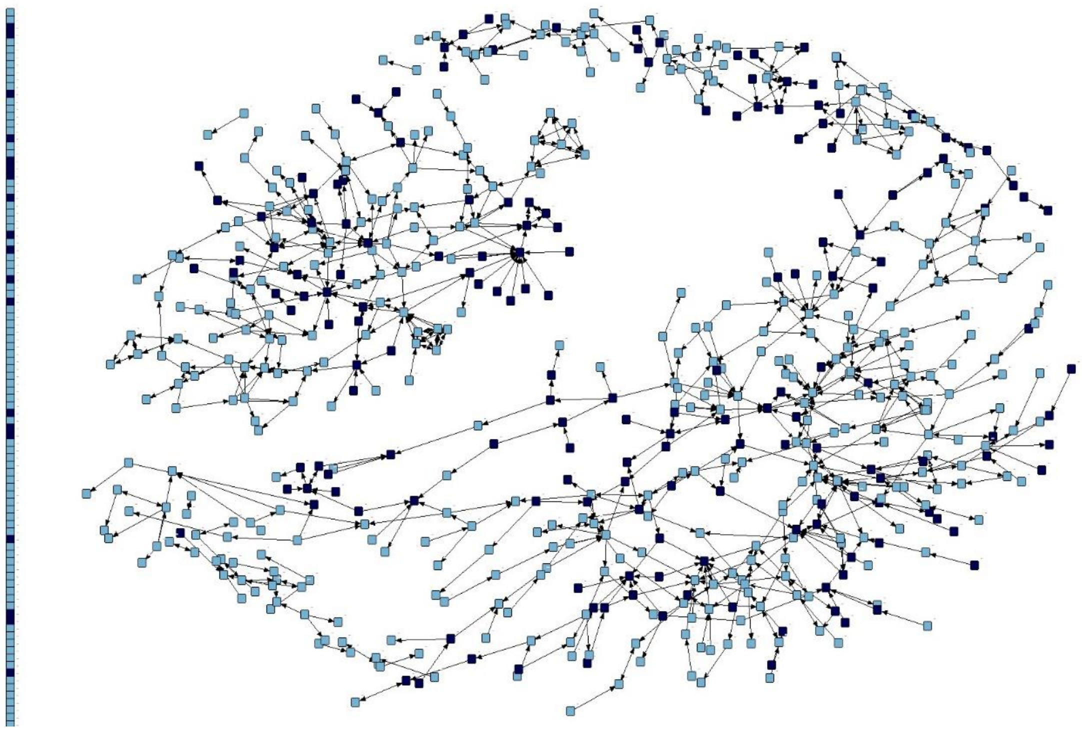


Figure A2: Graphical representation of the homeless friendship network



Notes: Sample of respondents living in the street, shelters or slums in January 2008. The graph is based on a spring embedding algorithm from UCINET. Each node represents one homeless. Lines between nodes indicate friendship relationship. Node colour denotes imprisonment, with dark blue describing individuals who have been to prison at least once and light blue describing individuals with non-criminal records. Source: Milan Homeless Survey, 2008

Table A1: Reduced form regressions and exclusion restriction

Panel A	(1)	(2)
Dependent variable=	*1 if prison during homelessness*	
Share of rainy days in one's first year as homeless	-0.368***	
	(0.119)	
Former inmates in one's first year as homeless[a)]		0.068***
		(0.016)
Panel B		
Dependent variable =	*Share of rainy days in one's first year as homeless*	*Former inmates in one's first year as homeless*
$Prison^{Before}$	-0.012	-16.063
	(0.009)	(132.099)
Observations	883	485

Notes: OLS regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the street of interview level. Constant not displayed. Sample of respondents living in the street, emergency shelters or slums in Janaury 2008. Estimates include dummies for the place of interview and for the month of entry. a) Coefficients and standard errors multiplied by 1000.

Table A2: First stage regression, by length of the homelessness spell

	Observations	*Proportion of the sample*	*Share of rainy days*	*Distribution of compliers*
	(1)	(2)	(3)	(4)
Panel A				
Short duration	433	0.5	5.118***	0.52
			(0.627)	
Long duration	450	0.5	3.579***	0.36
			(1.290)	
All sample	883	1	4.970***	
			(0.526)	
Panel B				
	Observations	*Proportion of the sample*	*Former inmates per month in one's first year as homeless* [a)]	*Distribution of compliers*
Short duration	308	0.5	0.079***	0.47
			(0.042)	
Long duration	181	0.5	0.058***	0.35
			(0.018)	
All sample	489	1	0.084***	
			(0.015)	

Notes: * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the street of interview level. Constant not displayed. Sample of respondents living in the street, in emergency shelters or in slums in January 2008. Variables are defined in the footnote of table 2. Estimates include dummies for the place of interview and for the month of entry. First stage linear coefficients for the sample of respondents with a length of the homelessness spell below ("short duration") and above the mean ("long duration"). Column 3, panel A report the coefficients for the first stage where the dependent variable is the number of best friends. Column 3, panel B reports first stage coefficients of a regression where the dependent variable is the share of friends in prison before homelessness. a) Coefficients and standard errors multiplied by 1000.

Table A3: Probability of imprisonment, by friends' homeless duration

Dependent variable =	*1 if prison during homelessness*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of "old" friends	-0.033***	-0.027**	-0.025*	-0.025*				
	(0.011)	(0.013)	(0.014)	(0.014)				
Share of "old" friends in prison before homelessness		0.621***	0.583***	0.584***				
		(0.066)	(0.067)	(0.068)				
Number of "new" friends					-0.023	-0.005	-0.058*	-0.058*
					(0.021)	(0.026)	(0.029)	(0.030)
Share of "new" friends in prison before homelessness						0.346***	0.316***	0.308***
						(0.089)	(0.098)	(0.097)
Duration < 1 year			0.046	0.059			---	---
			(0.057)	(0.077)			---	---
Duration			-0.004	-0.003			-0.005	-0.008
			(0.003)	(0.004)			(0.006)	(0.007)
Duration squared [a)]				-0.006				0.015
				(0.019)				(0.019)
Controls	no	no	yes	yes	No	no	yes	yes
Place of interview dummies	yes	yes	yes	yes	Yes	yes	yes	yes
Month of entry dummies	yes	yes	yes	yes	Yes	yes	yes	yes
Friends characteristics	no	no	yes	yes	No	no	no	no
Economic indicators	no	no	yes	yes	No	no	no	no
Observations	883	397	385	385	883	203	165	165

Notes: OLS regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the street of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Controls include a dummy equal to one if respondent has been in prison before homelessness, age, age squared, female, 2 dummies for educational level (primary education and secondary education and no education is the omitted category), three dummies equal to 1 if the respondent is Italian, Moroccan or Romanian; Place of interview is equal to 1 for homeless interviewed in the street, equal to 2 for homeless interviewed in shelters and equal to 3 for homeless interviewed in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014).

Table A4: Probability of imprisonment, by mutual relationship

Dependent variable =	*1 if prison during homelessness*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of reciprocal friends	-0.020	-0.024	-0.033	-0.033				
	(0.020)	(0.024)	(0.027)	(0.026)				
Share of reciprocal friends in prison before homelessness		0.337***	0.280**	0.281**				
		(0.128)	(0.113)	(0.114)				
Number of non-reciprocal friends					-0.042***	-0.035**	-0.027**	-0.027**
					(0.008)	(0.013)	(0.012)	(0.012)
Share of non-reciprocal friends in prison before homelessness						0.708***	0.718***	0.716***
						(0.067)	(0.067)	(0.065)
Duration < 1 year			-0.043	-0.028			0.069*	0.057
			(0.093)	(0.131)			(0.038)	(0.055)
Duration			-0.006	-0.005			-0.003	-0.003
			(0.004)	(0.006)			(0.002)	(0.003)
Duration squared [a)]				-0.000				0.000
				(0.000)				(0.000)
Controls	no	no	yes	yes	no	no	yes	Yes
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes	Yes
Month of entry dummies	yes	yes	yes	yes	yes	yes	yes	Yes
Friends characteristics	no	no	yes	yes	no	no	no	No
Economic indicators	no	no	yes	yes	no	no	no	No
Observations	883	202	182	182	883	453	409	409

Notes: OLS regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the street of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Controls include a dummy equal to one if respondent has been in prison before homelessness, age, age squared, female, 2 dummies for educational level (primary education and secondary education and no education is the omitted category), three dummies equal to 1 if the respondent is Italian, Moroccan or Romanian; Place of interview is equal to 1 for homeless interviewed in the street, equal to 2 for homeless interviewed in shelters and equal to 3 for homeless interviewed in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014).

Table A5: Probability of imprisonment by network size

Dependent variable =	*1 if prison during homelessness*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of friends till the 2° degree	-0.006	-0.002	-0.004*	-0.004*				
	(0.004)	(0.002)	(0.002)	(0.002)				
Share of friends till the 2° degree in prison before homelessness		0.997***	0.950***	0.947***				
		(0.117)	(0.112)	(0.110)				
Number of friends till the 3° degree					-0.003	-0.000	-0.001	-0.001
					(0.002)	(0.002)	(0.002)	(0.002)
Share of friends till the 3° degree in prison before homelessness						1.030***	0.991***	0.989***
						(0.125)	(0.120)	(0.119)
Duration < 1 year			0.003	-0.014			0.006	-0.005
			(0.074)	(0.094)			(0.073)	(0.091)
Duration			-0.005	-0.006*			-0.005	-0.006
			(0.003)	(0.004)			(0.003)	(0.004)
Duration squared [a)]				0.000				0.000
				(0.000)				(0.000)
Controls	no	no	yes	yes	no	no	yes	Yes
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes	Yes
Month of entry dummies	yes	yes	yes	yes	yes	yes	yes	Yes
Friends characteristics	no	no	yes	yes	no	no	no	No
Economic indicators	no	no	yes	yes	no	no	no	No
Observations	883	552	500	502	883	552	500	500

Notes: OLS regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the street of interview level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. a) Coefficients and standard errors multiplied by 1000. Place of interview is equal to 1 for homeless interviewed in the street, equal to 2 for homeless interviewed in shelters and equal to 3 for homeless interviewed in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014).

Table A6: Probability of imprisonment

Dependent variable =	1 if prison during homelessness						
	IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of best friends	-0.111***	-0.115***	-0.130***	-0.129***	-0.120***	-0.104**	-0.207**
	(0.028)	(0.032)	(0.038)	(0.037)	(0.044)	(0.050)	(0.101)
Share of friends in prison		0.580***	0.582***	0.559***	0.726***	0.838***	0.863**
before homelessness		(0.216)	(0.180)	(0.182)	(0.192)	(0.167)	(0.435)
Duration < 1 year			0.055	0.018	0.064	0.138	0.124
			(0.038)	(0.054)	(0.053)	(0.096)	(0.174)
Duration			0.000	-0.001	0.001	0.007	0.046
			(0.000)	(0.002)	(0.004)	(0.005)	(0.096)
Duration squared [a)]				0.010	-0.004	-0.095	-0.016
				(0.012)	(0.011)	(0.087)	(0.065)
Duration cubic [a)]						0.000	
						(0.000)	
Controls	no	no	yes	yes	yes	yes	yes
Place of interview dummies	yes	yes	yes	yes	yes	yes	yes
Month of entry dummies	no	no	yes	yes	yes	yes	yes
Friends characteristics	no	no	no	no	yes	yes	yes
Economic indicators	no	no	no	no	yes	yes	yes
Year of entry dummies	no	no	no	no	no	no	yes
Observations	883	483	478	478	477	477	477

Notes: Linear regressions. * denotes significance at 10 percent level; ** at 5 percent level; *** at 1 percent level. Robust standard errors in parenthesis clustered at the month by year level. Constant not displayed. Sample of homeless living in the street, emergency shelters or slums in January 2008. Variables are defined in footnote of table 2. a) Coefficients and standard errors multiplied by 1000. Controls include a dummy equal to one if respondent has been in prison before homelessness, age, age squared, female, 2 dummies for educational level (primary education and secondary education and no education is the omitted category), three dummies equal to 1 if the respondent is Italian, Moroccan or Romanian; Place of interview is equal to 1 for homeless interviewed in the street, equal to 2 for homeless interviewed in shelters and equal to 3 for homeless interviewed in slums; Friends' characteristics include: average years of education, average duration of the homeless spell and the share of friends with the same nationality as the one of the respondent. Economic indicators include the total GDP in level (at market price) by quarter and the number of unemployed (those looking for a job older than 15 years old) by quarter in the region of Lombardy between 1993 and 2008 (ISTAT, 2014). The number of best friends is instrumented with the share of rainy days during one's first year as an homeless person and the share of friends in prison before homelessness is instrumented with the fraction of former inmates during one' first year as an homeless person.