de Graaff, Thomas; Rouwendal, Jan; Bernasco, Wim; Steenbeek, Wouter

Conference Paper

Social interaction and the spatial concentration of criminality

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:
European Regional Science Association (ERSA)

Social interaction and the spatial concentration of criminality

Wim Bernasco[a] Thomas de Graaff[b] Jan Rouwendal[b*] Wouter Steenbeek[a]

[a] *Netherlands Institute for the Study of Crime and Law Enforcement*
VU University, Amsterdam, The Netherlands

[b] *Department of Spatial Economics, Faculty of Economics and Business Adminstration*
VU University, Amsterdam, The Netherlands

February 17, 2012

[DRAFT VERSION; DO NOT CITE]

Abstract

It has often been observed that there is substantial spatial variation in criminality, i.e. criminality clusters in neighborhoods. Differences in neighborhood characteristics are one possible reason, social interactions another. In this paper we use detailed data on the residential location of criminals to disentangle the effects of individual characteristics, neighborhood characteristics and social interaction on criminality. Our basic model is an individual binomial logit model for the probability of being a criminal which we use to extract neighborhood effects. In a second stage, we model neighborhood effects, where we use as explanatory variables physical and social neighborhood characteristics such as characteristics of the housing stock and the demographic composition. We also include the share of criminals to be able to measure a social interaction effect. Note that this approach takes into account unobserved neighborhood characteristics. Since these may affect the criminality rate we instrument for this variable in a two stage estimation procedure.

Keywords: social interactions, neighborhoods, crime
JEL-classification: R1, R2

*Corresponding author. Tel +31 (0)20 5986090. Email: jrouwendal@feweb.vu.nl. We would like to thank the participants of the Vth Congress of the Spatial Econometric Association (SEA) in Toulouse 2011 for useful comments. This paper is still work in progress.

1 Introduction

Economists have traditionally focused on market interactions, but in the second half of the 20th century many economists began to see their discipline as relevant for all kinds of resource allocation processes including that of crime (see Manski 2000). Within the economic literature on crime the importance of social interactions has frequently been emphasized. An important reason for conjecturing such a relationship is the large amount of spatial variation in crime rates (see, e.g., Glaeser et al. 1996). The mechanism of social interaction is not always made completely explicit, but the general idea is clear enough: if one's peer group contains a relatively large numbers of criminals, the probability to become a criminal increases.

One question that has received relatively little attention in this literature is the spatial level at which the relevant type of social interaction takes place. For instance, Glaeser et al. (1996) use data that refer to metropolitan areas, but is seems probable that the relevant spatial level is much smaller. A priori, it seems much more likely that neighborhoods are the relevant level. Moreover, attention is often concentrated on the location where criminal acts are carried out, whereas residential locations of the criminals seem at least as relevant to study.

In this paper we use a rich dataset on the residential location of criminals to investigate the relationship between criminality, neighborhood composition and its associated social interaction. More specifically, we use the 4-digit postal code classification of the Netherlands as the relevant size of the neighborhoods. Although this definition is partly the result of practical considerations, the Dutch postal codes also have several substantive advantages. First, the boundaries of the postal codes are constructed so that there are few or no mobility constraints within the neighborhood. Moreover, the postal codes classification refers to 'neighborhoods' in cities, but encompass small villages in its entirety. Thus, its geographical area is usually easily walkable, with meaningful geographic and social boundaries.

Our basic research question boils down to the question whether an individual's decision to commit crime depend on whether other neighborhood residents commit crime? In this case, the neighborhood residents are considered as his peers. We aim to do so by disentangling the effects of individual characteristics, neighborhood characteristics and social interaction on criminality. Our methodological approach is as follows. First, we estimate for each neighborhood an average probability level of becoming a criminal by using a large individual data of being a criminal or not using sex and age as co-variates. In the second stage, we regress various neighborhood characteristics, including a social interaction effect, on the neighborhood fixed effects using an instrumental variable approach. As instruments we use a spatial and a social peer reference group (*cf.* Walker et al. (2011)).

The paper is structured as follows. The next section discusses the literature concerning criminality and social interactions mainly from a (very rich) criminality literature. The subsequent section deals with the methodological approach. Section 4 treats extensively the data we use.

Section 5 offers the results where after the last section concludes concisely.

2 Literature

In criminology, it has long been observed that peer delinquency and individual delinquency are correlated.[1] Two mechanisms are hypothesized to underlie that correlation. Differential Association Theory (Sutherland 1947) argues that criminal behavior is learned from delinquent peers; this includes learning techniques of committing delinquent behavior, as well as learning motives and attitudes that promote delinquent behavior, i.e. techniques of neutralization (Sykes and Matza 1957). Further elaborated by Akers (1985), it is argued that individuals are *influenced* by delinquent peers through processes such as social reinforcement and imitation. The general proposition of these theories is that the excess of definitions favorable to deviance over definitions unfavorable to deviance enhances the probability of offending.

A contrasting view argues, firstly, that delinquent behavior is caused by other factors, such as weak social bonds or low self-control (Gottfredson and Hirschi 1990), and secondly, that delinquents or criminals associate with each other precisely because they choose to be friends with others who are similar to themselves. Thus, in this view *selection* processes lead to social networks of delinquents. The selection of friends need not be solely based on behavioral similarities, but also on proximity. People located near each other are more likely to become friends with each other than with more distant others. Such proximity selection (Festinger et al. 1950) may also be partly responsible for why delinquents befriend delinquent peers. For example, if *delinquents cluster within neighborhoods and if friend selection is based partly on proximity*, then each delinquent is more likely to befriend other delinquents rather than non-delinquents – a 'decision' process independent of behavioral similarity. Selection and influence are depicted in abstract form in Figure 1.

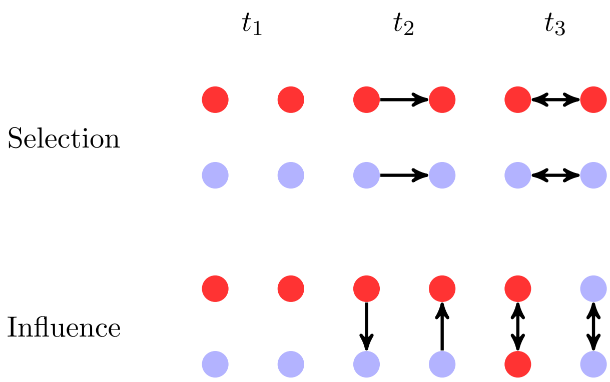

Figure 1 – Abstract representation of selection and influence processes

[1]Here, delinguency is defined as criminality amongst youngsters.

The correlation between peer delinquency and individual delinquency is thus hypothesized to be affected by processes of selection and influence in social interactions. Manski (2000) argues that a rigorous identification of social interactions will require clear thinking and adequate data. For the latter, usually experimental data are more preferred than outcome data. There are a few studies that go a long way to do precisely this. For example, Young et al. (2011) measured the complete peer networks in school classes to establish whether individuals' self-reported delinquency matched peers' ratings of their delinquency. However, establishing complete social networks of individuals is a difficult, invasive, and thus very costly challenge. Moreover, in one of the first studies to employ longitudinal network analyses to study the causal ordering of selection and influence, Weerman (2011) shows that only the average delinquency level of someone's friends in the school network has a significant, although relatively small, effect on individual delinquent behavior. Social influence thus seems to be most important to explain the correlation between peer delinquency and individual delinquency.

There are, in addition, arguments against using the full extent of social networks to estimate social interactions. First, the composition of social networks that includes friends and other relations that are the object of choice is likely to be endogenous, as people may have incentives to select as peers others who are as delinquent as they are themselves (e.g. two criminals cannot betray each other as each knows that such behavior will be reciprocated by the other). Parents, siblings, teachers and neighborhood residents may be more exogenous. Second, social interactions include mechanisms that do not rely on the identification of other individuals. For example, an individual's decision to commit crime may be affected by merely observing the behavior of unknown others, or even by just observing the outcomes of it, and inferring the behavior.

In this paper, we focus on the *social influence* that *neighborhood peers* exert on *individual delinquency*. Our main hypothesis corresponds to Weerman (2011):

> "...a high (mean) level of delinquency among peers increases the chances that less delinquent adolescents adapt their behavior to that of their friends." (p. 257)

Thus, we expect that one's behavior is influenced by observing or 'having knowledge' about the behavior of others. Relevant examples for the purposes of this paper are (*i*) see crime take place, (*ii*) hear about crime taking place from offenders and/or victims in one's peer group (*iii*) see the results of crime (*iv*) become a victim of crime (*v*) be subjected to specific 'norms' about crime. We test this hypothesis using a different method than longitudinal social network analysis (see next section).

Our focus on social influence in the neighborhood instead of the family or school is motivated by the wealth of research with regard to neighborhood influences on individual delinquency (for an overview of not only crime-related outcomes, see Sampson et al. (2002)). As Shaw and McKay (1969) pointed out in their highly influential work:

> "Heavy concentration of delinquency in certain areas means [...] that boys living

in these areas are in contact not only with individuals who engage in proscribed activity but also with groups which sanction such behavior and exert pressure upon their members to conform to group standards. [...] In contrast with the areas of concentrations of delinquents, there are many other communities where the cases are so widely dispersed that the chances of a boy's having intimate contact with other delinquents or with delinquent groups is comparatively slight." (p. 174)

Whereas subsequent scholars have often focused on the processes of social disorganization, i.e. social cohesion and informal social control within neighborhoods (Sampson et al. 1997), the influence of delinquent neighbors as well as the choice-constraining effect of neighborhoods for friend selection has been neglected. Whereas current criminological studies with regard to selection and influence often investigate peers in general or specifically adolescent social networks within schools, we argue that neighborhoods are important contexts for the current discussion.

3 The model

This sections presents the methodology. It first lays out the basic model along the lines of Walker et al. (2011). Subsequently, it deals with the issue of correct identification of the social interaction effect. The last subsection addresses the problem that our social interaction effects, the number of criminals within a neighborhood, is endogeneous.

3.1 Introduction

The model we use focuses on the choice to be a criminal. Our basic model is binomial: one can choose to be a criminal, or not. The choice depends on personal and neighborhood characteristics, not all of which are observed. Let C_{ij} be a zero-one variable that indicates whether individual i in neighborhood j is a criminal. The probability that C_{ij} equals 1 (indicating that this person is a criminal) depends on personal characteristics X_i, and on neighborhood characteristics Z_j. A social interaction effect is present if the expected value of the variable C_{ij} in neighborhood j has an impact on the probability that a particular individual i is a criminal. Since we are not informed about all the relevant characteristics, we introduce two random variables representing unobserved characteristics: ϵ_i for unobserved personal characteristics and ξ_j for unobserved neighborhood characteristics. We now define a latent variable y_{ij} that is linear in these characteristics:[2]

$$y_{ij} = \beta X_i + \gamma Z_j + \delta E(C_j) + \xi_j + \epsilon_i. \tag{1}$$

When this latent variable (y_{ij}) takes on a positive value, then $C_{ij} = 1$, otherwise $C_{ij} = 0$.

If we assume the random variable ϵ_i to be extreme value type I distributed the probability

[2] A simple extension would be to allow for cross effects by introducing the terms $\theta X_i Z_j$.

that $C_{ij} = 1$ is given by the logit expression:

$$\Pr(C_{ij} = 1) = \frac{e^{\beta X_i + \gamma Z_j + \delta E(C_j) + \xi_j + \epsilon_i}}{1 + e^{\beta X_i + \gamma Z_j + \delta E(C_j) + \xi_j + \epsilon_i}} \tag{2}$$

Without the social interaction and unobserved neighborhood effects (i.e., $\delta = \xi_j = 0$), this is a standard binomial logit model. When there is social interaction, but no unobserved heterogeneity ($\xi_j = 0$), this is the logit version of the binomial model of Brock and Durlauf (2001).

The unobserved heterogeneity term ξ_j captures neighborhood characteristics that may have an impact on the individual's probability to become a criminal, but are unobserved by the analyst. The importance of such unobserved heterogeneity in discrete choice models has been first analyzed thoroughly by Berry et al. (1995) in their seminal study of the automobile market. Their approach has been used in other fields as well. For instance, Walker et al. (2011) have used a model like (2), but without neighborhood variables Z_j, to study the effect of social interaction on traffic mode choice.

3.2 Identification

Berry et al. (1995) suggest a two-stage procedure. In the first step the neighborhood-specific terms are taken together in a single neighborhood constant α_j, where the probability of becoming a criminal is now given by:

$$\Pr(C_{ij} = 1) = \frac{e^{\beta X_i + \alpha_j}}{1 + e^{\beta X_i + \alpha_j}}, \tag{3}$$

which can be estimated in the usual way. α_j can then be defined as a neighborhood specific attractivity index to become a criminal, conditional on as much individual characteristics, X_i, as possible. Note that this is a general measure, and is still composed of various exogeneous, endogeneous and contextual effects which can affect the probability to become a criminal on a neighborhood level Manski (1993).

In the second stage the alternative specific constants are analyzed further by writing them again as:

$$\alpha_j = \gamma Z_j + \delta E(C_j) + \xi_j \tag{4}$$

and using techniques for linear equations.

The unobserved heterogeneity terms ξ_j are now the residuals of the linear regression equation. A complication is that OLS cannot be used, since $E(C_j)$ must be expected to be correlated with ξ_j. The reason is that a high value of ξ_j makes it more likely that a particular individual in the neighborhood is a criminal, which tends to increase $E(C_j)$. Hence the error term is not independent of the explanatory variables. In the next subsection we will propose a solution to this problem using an instrumental variable approach.

Manski (1993) looked at identification issues within a linear model with social interactions in which there are endogenous interaction effects as well as contextual effects. In our model the variable $E(C_j)$ embodies an endogenous social interaction effect, while contextual effects may be included in the vector Z_j when it contains variables like the average age of neighborhood inhabitants. In Manski's model, the two effects cannot be distinguished. Brock and Durlauf (2001) have shown that the nonlinearity that occurs in a discrete choice model like (2) has identifying power. They develop a set of conditions under which all the remaining parameters are identified. These conditions apply to the model (2) when the term referring to unobserved heterogeneity is absent.

The model (2) is identified if the parameters β and α in (3) are identified and if the parameters γ and δ in (4) are identified. The first is not a problem (see Manski 1988). The estimated xi_j's are complex nonlinear functions of the variables C_{ij} and X_i, whereas in Manski's linear model C_{ij} is on the left-hand side of the equation. This is the reason why Manski's identification problem does not occur in the context of a binomial choice model. However, there is another problem that has to be faced: the term ξ_j, which represents unobserved heterogeneity has an impact on all C_{ij}'s and therefore also on $E(C_j)$. The implication is that $E(C_j)$ is potentially correlated with ξ_j. In the next subsection we will propose an instrumental variable strategy to solve this problem.

3.3 Endogeneity

We need additional variables that have no direct impact on α_j, are correlated with $E(C_j)$, but not correlated with ξ_j. We adopt the approach by Walker et al. (2011) and define two types of instruments. A *spatial* reference group, or the average social interaction effect of the adjacent postal codes and a *social* reference group, variables that indicate whether inhabitants of a neighborhood share similar socio-economic characteristics.[3]

The intuition behind this approach is – theoretically at least – straightforward. $E(C_j)$ is defined as the expected number of criminals within a neighborhood. This can also be seen an the probability to encounter a criminal within neighborhood j. Note that one of the main assumptions of our model is that social interactions take place within a neighborhood. Thus, spatially lagged encounter probabilities $WE(C_j)$ are not correlated with the neighborhood specific effect, but might contain information about $E(C_j)$.

Similarly, the social distance in the neighborhood to other groups should be correlated with $E(C_j)$ (it contains information about the strength of the network and thus the intensity of the social interactions) and it can be argued that the correlation with α_j is rather weak. For instance, it has not been proven that a neighborhood's age structure correlates with the propensity of becoming a criminal other than via group interactions.

Section 5 deals further with this implementations and treats the instruments more in detail.

[3] Actually, Bayer et al. (2004) have suggested that an instrument can be computed by computing the 'equilibrium' values for $E(C_j)$ that would obtains if all ξ_j's would be equal to 0. This computed variable is certainly not correlated with the ξ_j's and almost certainly strongly correlated with the $E(C_j)$'s. We elaborate further on this approach in Appendix A and leave the calculation of this instrument for near further research.

4 Data

Criminality is notoriously difficult to measure. Because criminal behavior is morally objectionable and legally sanctioned, most people are unwilling to confess their involvement in crime, either to law enforcement or to researchers. While there are surveys asking adolescent subjects for their involvement in crimes and rule breaking (see, amonst others, Farrington et al. 1996, Steffensmeier and Allan 1996, Piquero et al. 2002), crime self-report surveys are rare amongst adult populations (Morselli and Tremblay 2004).

To measure criminality, we used anonymized national population data from the Dutch National Police. The police information system from which the data were extracted, contains data on all individuals that have been arrested by the Dutch police as criminal suspects (the large majority are subsequently convicted, some obtain a 'transaction' from the prosecutor's office, and a few are dismissed). The data contain some personal characteristics (sex, age, nationality, country of birth, postal code of residential address) and also contain details about all crimes of which the individual has been suspected (including the dates and the types of crime). In the analysis in this paper we use involvement in any crime(s) in the year 2006 as the dependent variable as well as involvement in (*i*) violent and (*ii*) property crime.

Because the police information system is used for investigative purposes, it is updated continuously, and updates include changes of address as well as removal of individuals after an expiration period, the length of which depends on the seriousness of their criminal record. The database used in this analysis was an archival copy of the information system included crimes already removed from the real 'living' information system. Data from special investigative services are excluded, so that tax and other economic crimes, social security fraud, and environmental crimes are underrepresented.

There are some disadvantages to using police records to measure criminality. First, a substantial percentage of crimes never comes to the attention of the police, either because there is not an individual victim to report it (e.g. drug dealing) or because the victim does not report the crime to the police (Goudriaan et al. 2004). Second, in most jurisdictions the police solve only approximately 20 percent of all crimes (Dodd et al. 2004). As a consequence, any estimate of criminality based on police data must be a severe underrepresentation. Third, specific surveillance or investigative strategies used by the police may result in some areas being more intensely supervised and investigated than others, resulting in an overrepresentation of these areas in the data. Fourth, police records have data on suspects and individuals charged with criminal offences, but some of these people may be unjustly suspected and will not be convicted subsequently in court. Notwithstanding these limitations police records seem to be the best available large-scale measures of criminality that we have.

To obtain a full population data set on criminal involvement in 2006 in The Netherlands, we used population data from Statistics Netherlands. The relevant table applies to January

1st, 2006 and cross-tabulates neighborhood of residence (4,028 neighborhoods) with age (20 categories, each 5 years width) and sex (male versus female). As the police records contain these three variables as well, both sources were combined to create a national dataset containing approximately 16 million persons with the following four variables:

1. neighborhood (4028 neighborhoods);
2. sex (male or female);
3. age (20 categories: 0–4 years, 5–9 years, 10–14 years, 15–19 years, etc.);
4. involved in crime in 2006 (yes or no).

Because in The Netherlands only individuals of age 12 and older can be prosecuted, age categories 0–4 years and 5–9 years were removed from the analysis. Ages 10–11 are included because the population data are available only in 5-years age categories. Because no individuals above age 89 were prosecuted in 2006, ages 90 and above were also removed from the analysis. The remaining dataset contains 14,298,733 individuals aged 10–89 in 2006.

Table 1 – Dutch population ages 10–89, January 1st, 2007. Absolute population size (#) and percentage involved in crime in 2006 (%), by sex and age category (source: police force Haaglanden Statistics Netherlands

	Female		Male	
	Population #	Criminal %	Population #	Criminal %
10–14 years	480,980	0.32	504,455	1.02
15–19 years	487,747	1.31	510,485	6.27
20–24 years	477,712	1.06	488,167	6.21
25–29 years	494,210	0.78	494,942	4.17
30–34 years	534,001	0.64	533,858	3.23
35–39 years	641,060	0.64	653,601	2.82
40–44 years	646,021	0.59	663,348	2.50
45–49 years	613,101	0.50	621,996	2.02
50–54 years	562,241	0.36	569,603	1.53
55–59 years	550,147	0.26	560,466	1.13
60–64 years	460,035	0.19	464,104	0.86
65–69 years	361,255	0.12	345,704	0.62
70–74 years	314,045	0.09	270,665	0.41
75–79 years	274,375	0.07	200,438	0.28
80–84 years	216,185	0.04	122,983	0.20
85–89 years	126,630	0.02	54,173	0.18
Total	7,239,745	0.50	7,058,988	2.49

For this population, Table 1 lists the percentages of criminal involvement in the year 2006 by sex and age category. The table confirms two stylized facts about criminality: the criminal involvement of men is five times larger than women's involvement (*cf.* Steffensmeier and Allan 1996, Mears et al. 1998), and criminal involvement of both sexes peaks during adolescence and early adulthood at ages 15–24 (*cf.* Blokland et al. 2005). On average, 1.5 percent of the 10–89 population became a crime suspect in 2006. For boys in the age category 15–24 years, the percentage is more than four times larger than the average.

The focus of our investigation is the percent of people in the neighborhood involved in crime.[4] "Spatial reference group" are neighborhoods. Substantive arguments for neighborhood as a valid peer group were already given in the literature overview. There are also several methodological arguments in favor of the neighborhood (instead of a larger or smaller areal unit): (1) smaller areas result in very skewed crime distributions (2) no areal data on smaller areas. Note: we use pc4 postal code areas as our definition of neighborhood. In short, following Walker et al. (2011), we assume that "these postal code boundaries delineate spatial peers and that individuals within a postal code are more similar, exerting a stronger influence than individuals who live outside of one's postal code".

The police records include the six-digit postal codes of the residential addresses of the individuals. Throughout the Netherlands there are about 435,000 six-digit postal code areas. In non-rural areas they are roughly the size of a football field and contain approximately 20 residential properties and 40 residents. As they were created with pedestrian postal delivery services in mind, single codes are nearly always on the same street, apply to adjacent properties, and is not subdivided by physical barriers that impede pedestrian or car transportation. In line with definitions of 'neighborhood' as a locus of social interaction elsewhere in the literature, our analysis uses the four-digit Dutch postal code number as the spatial unit of analysis, i.e. a spatial aggregation of the six-digit postal code. Many other studies in The Netherlands have used the four-digit postal code as a neighborhood delineation criterion (Van Wilsem et al. 2006, Nieuwbeerta et al. 2008, Walker et al. 2011).

Geographically, the Netherlands is a small country with a total land surface of 41,526 square kilometers. The total country consists of 4,028 four-digit postal code areas with an average surface of 10.31 square kilometer and an average population of 4,073 inhabitants. Similar to US census tracts, the sizes of these 'neighborhoods' depend on population density. In urban areas where population densities are high, the surface of neighborhoods tend to be relatively small, while they are larger in rural areas where population densities are low.

[4] When the peer group is the neighborhood, the chance of interaction with a criminal is affected not only by the relative number of criminals, but also by the size of the area. Thus, the social interaction effect can alternatively be defined as the percent of people per square mile exhibiting a given behavior.

5 Results

In addition to the fact whether someone is a criminal or not, we also have information about the type of crime. We therefore choose to look at two types of crime as well: violent crime, such as assault and domestic violence, and property crime, such as burglary, shoplifting and vandalism. Because violent crime has a strong reciprocal nature (assault often takes place for reasons of revenge), we hypothesize that the social interaction effect for violent crime is larger than for property crime.

Table 2 presents the estimation results of equation (3). The socio-demographic variabkles included are a sex indicator (0 for males, 1 for females), age (measured categorically as 10–14$=-1$, 15–19$=0$, 20–24$=1$, $\ldots$, i.e. centered on the peak of the age-crime curve) and age squared. Of the 4011 neighborhoods there were 290 in which not a single resident offended in 2006, making it impossible to estimate a neighborhood specific constant term for the general model. In 323 neighborhoods not a single violent act of crime took place and property crime was absent in 971 neighborhoods.[5]

Estimation results confirm that males are much more likely to become criminals than females and that the propensity to become a criminal first increases with age and then decreases. Violent crime behave more or less similarly as crime in general and property crime is caused by much younger criminals.

Table 2 – Choice models (log-odds of choice to become a criminal—$\Pr(C_{ij}=1)$)

	All crime		Violent crime		Property crime	
Parameter	Estimation	S.E.	Estimation	S.E.	Estimation	S.E.
Female	-1.592	0.006	-1.830	0.007	-1.092	0.010
Age	0.049	0.002	0.079	0.002	-0.048	0.004
Age^2	-0.025	0.000	-0.028	0.000	-0.019	0.000
# Observations	14,221,511		14,206,517		13,812,878	
# Parameters	3 + 3721 constants		3 + 3688 constants		3 + 3040 constants	
Log-likelihood	$-983,941.11$		$-825,950.55$		$-338,008.36$	

A higher value of a neighborhood-specific constant means that inhabitants of the neighborhood are genericly more likely to be suspected of criminal involvement. Because of its individual nature, the first-stage model is silent about the reasons for this situation: they have to be sorted out in the second stage. It is clearly not insightful to report all the values of the estimates for α. The kernel density estimates in Figure 2 reveal that all three density functions are single peaked and almost symmetric. Again the kernel density estimate of violent crime is remarkably similar

[5]Usually, only the smallest neighborhoods with few of no criminals fall out of the estimation, which might invoke a selection bias. Note, however, that the number of observations decrease much slower than the number of neighborhoods.

to that of crime in general, where the kernel density estimate of property crime clearly has a higher mean and standard error.

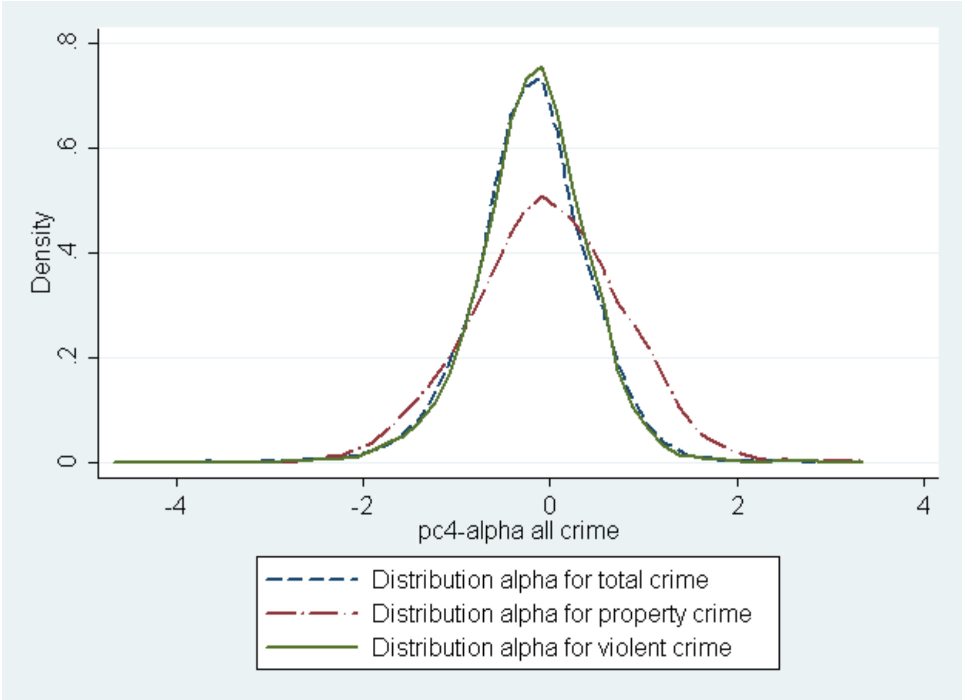

Figure 2 – Kernel density estimates of the distribution of α

Results of the analysis of the neighborhood specific constants are reported in Tables 3 and 4. The first table reports the results of the first stage regression on our social interaction variable (*% involved in crime*). The second table report the second stage results of the regression on the α's. As instruments of *% involved in crime* we have chose to include only the spatially lagged variable (thus, $W \times$*% involved in crime*) and the neighborhood age structure (as a social lag, *cf.* Walker et al. (2011)).

As Table 3 clearly shows, the instruments we use are relevant. The spatial lag (W *% involved in crime*) is very significant and shows a positive correlation with *% involved in crime* as expected. The age structure, though significant, is more difficult to interpret directly. In theory, these estimates should give us an imputed social interaction variable *% involved in crime* exogeneous to our vector of α's. The results of this second stage estimation are reported in Table 4.

Obviously, our field variable *% involved in crime* is positive and significant – both statistically and impact wise. One additional percent of criminals within a neighborhood increases α with about 0.3, which entails that the probability to become a criminal increases with 35%. Interestingly, this field effect is highest for property crime vis-á-vis violent crime, although we hypothesized that, because of its recriprocal nature, violent crime should be more susceptible to social interaction effects. Property crime is, however, associated with very young criminals (see, e.g., the results in Table 2) and perhaps they are more susceptible to social interactions

Table 3 – 2SLS estimation on α—first stage regression on *% involved in crime*

	All crime		Violent crime		Property crime	
Parameter	Estimation	S.E.	Estimation	S.E.	Estimation	S.E.
# Addresses per hectare	-0.001	0.001	-0.001	0.001	-0.004	0.001
% owner-occupied housing	-0.013	0.001	-0.013	0.001	-0.015	0.001
Average well-being	-0.114	0.035	-0.118	0.035	-0.095	0.042
W % involved in crime	0.284	0.020	0.283	0.020	0.289	0.023
% 0–5 years	2.326	1.607	2.337	1.621	7.123	2.072
% 5–10 years	4.379	1.790	4.375	1.791	11.237	2.481
% 10–15 years	-8.443	1.765	-9.005	1.770	-6.400	2.344
% 15–20 years	5.981	1.425	6.069	1.426	4.707	1.869
% 20–25 years	2.843	1.035	2.576	1.045	6.772	1.281
% 25–30 years	6.806	1.406	7.267	1.422	10.483	1.724
% 30–35 years	7.130	1.517	6.348	1.528	12.764	1.946
% 35–40 years	-7.964	1.483	-8.106	1.496	-11.969	1.940
% 40–45 years	8.503	1.431	8.371	1.451	14.391	1.842
% 45–50 years	1.933	1.499	1.602	1.496	8.231	2.153
% 50–55 years	4.625	1.460	4.209	1.468	6.609	1.934
% 55–60 years	2.448	1.202	2.867	1.194	8.376	1.354
% 60–65 years	-5.225	1.652	-5.892	1.661	-8.433	2.186
% 65–70 years	7.679	1.900	8.191	1.900	11.908	2.783
% 70–75 years	-4.706	2.149	-5.623	2.147	6.433	3.063
Intercept	0.355	0.494	0.498	0.494	-2.746	0.668
Instrument (ir)relevance (Shea test)	39.84	0.000	39.61	0.000	41.71	0.000
# Observations	3,709		3,678		3,035	
R^2	0.308		0.306		0.319	

than older criminals. Unfortunately, two of the three Durbin-Wu-Hausman reject exogeneity. Alternative specifications (with different instruments) show that, alhough all three tests never pass exogeneity simultaneously, our results regarding our social interaction variable are rather robust.

The remainder of the variables give a more or less coherent and intuitive appealing picture. A higher housing density, although very small, correlates positively with more crime. More owner-occupied housing correlates with less crime just as a larger average well-being within a neighborhood (measured as social-economic status).

Table 4 – 2SLS estimation on α—second stage

	All crime		Violent crime		Property crime	
Parameter	Estimation	S.E.	Estimation	S.E.	Estimation	S.E.
% Involved in crime	0.285	0.013	0.258	0.014	0.339	0.021
# Addresses per hectare	0.000	0.000	0.000	0.000	0.002	0.001
% Owner-occupied housing	-0.009	0.000	-0.008	0.000	-0.012	0.001
Average well-being	-0.026	0.011	-0.034	0.012	0.000	0.021
Intercept	0.114	0.058	0.098	0.061	0.308	0.101
Durbin-Wu-Hausman test	0.747	0.387	28.754	0.000	0.813	0.367
# Observations	3,709		3,678		3,035	
R^2	0.673		0.604		0.537	

6 Conclusion

It has often been observed that there is substantial spatial variation in criminality, i.e. criminality clusters in neighborhoods. Differences in neighborhood characteristics are one possible reason, social interactions another. The main aim of this paper was to disentangle the effects of individual characteristics, neighborhood characteristics and social interaction on criminality using a rich dataset that provides information about the residential location of criminals in the Netherlands. Our main (but very preliminary) results are that every percentage increase of criminals within a neighborhood increases the change on becoming a criminal with 35%. We still have to investigate further whether this results holds within a variety of alternative specifications, but, if so, then this is a very sizeable effect. Moreover, the impact of social interactions seem to depend on the type of crime as well. Property crime seems be more susceptible to social interactions than violent crime. Probably, this is due to the fact that property crime is much more associated with younger criminals than violent crime.

Obviously, the results are still very preliminary and there is much scope for further research. First, and most importantly, we have to check whether our instruments are valid and whether we can find alternative instruments, such as spatial weights matrices of criminals for each age group. Secondly, although our first stage logit estimation performs conform expectations we might still add additional information such that we remove most of the *individual* variation in the first binary logit stage. Thirdly, and finally, we need to look more in the impact variation of social interaction across types of crime just as in Glaeser et al. (1996), both as an indicator for the validity of our results and to see whether some types of crime are not susceptible to social interactions as well in the Netherlands.

References

Akers, R. L.: 1985, *Deviant Behavior: A Social Learning Approach*, California Wadsworth Publishing Company, Belmont.

Bayer, P., McMillan, R. and Rueben, K. S.: 2004, What drives racial segregation? New evidence using Census microdata, *Journal of Urban Economics* **56**, 514–535.

Berry, S., Levinsohn, J. and Pakes, A.: 1995, Automobile Prices in Market Equilibrium, *Econometrica* **63**, 841–890.

Blokland, A. A. J., Nagin, D. S. and Nieuwbeerta, P.: 2005, Life span offending trajectories of a dutch conviction cohort, *Criminology* **43**, 919–954.

Brock, W. A. and Durlauf, S. N.: 2001, Discrete choice with social interactions, *Review of Economic Studies* **68**, 235–260.

Dodd, T., Nichlas, S., Povey, D. and Walker, A.: 2004, Crime in england and wales 2003/2004. Home Office Statistical Bulletin No.10/04. London: Home Office.

Farrington, D. P., Loeber, R., StouthamerLoeber, M., VanKammen, W. B. and Schmidt, L.: 1996, Self-reported delinquency and a combined delinquency seriousness scale based on boys, mothers, and teachers: Concurrent and predictive validity for African-Americans and Caucasians, *Criminology* **34**, 493–517.

Festinger, L., Schachter, S. and Back, K.: 1950, *Social Pressures in Informal Groups: A Study of Human Factors in Housing*, Harper, New York.

Glaeser, E. L., Sacerdote, B. and Scheinkman, J. A.: 1996, Crime and social interactons, *Quarterly Journal of Economics* **111**, 507–548.

Gottfredson, M. R. and Hirschi, T.: 1990, *A general theory of crime*, Stanford University Press, Palo Alto.

Goudriaan, H., Lynch, J. P. and Nieuwbeerta, P.: 2004, Reporting to the police in Western nations: A theoretical analysis of the effects of social context, *Justice Quarterly* **21**, 933–969.

Manski, C. F.: 1988, Identification of Binary Response Models, *Journal of the American Statistical Association* **83**, 729–738.

Manski, C. F.: 1993, Identification of Endogenous Social Effects: The Reflection Problem, *Review of Economic Studies* **60**, 531–542.

Manski, C. F.: 2000, Economic analysis of social interactions, *Journal of Economic Perspectives* **14**, 115–136.

Mears, D. P., Ploeger, M. and Warr, M.: 1998, Explaining the gender gap in delinquency: Peer influence and moral evaluations of behavior, *Journal of Research in Crime and Delinquency* **35**, 251–266.

Morselli, C. and Tremblay, P.: 2004, Criminal achievement, offender networks and the benefits of low self-control, *Criminology* **42**, 773–804.

Nieuwbeerta, P., Mccall, P. L., Elffers, H. and Wittebrood, K.: 2008, Neighborhood characteristics and individual homicide risks - Effects of social cohesion, confidence in the police, and socioeconomic disadvantage, *Homicide Studies* **12**, 90–116.

Piquero, A. R., MacIntosh, R. and Hickman, M.: 2002, The validity of a self-reported delinquency scale - Comparisons across gender, age, race, and place of residence, *Sociological Methods & Research* **30**, 492–529.

Sampson, R. J., Morenoff, J. D. and Gannon-Rowley, T.: 2002, Assessing "neighborhood effects": Social processes and new directions in research, *Annual review of sociology* **28**, 443–478.

Sampson, R. J., Raudenbush, S. W. and Earls, F.: 1997, Neighborhoods and violent crime: A multilevel study of collective efficacy, *Science* **277**, 918–924.

Shaw, C. R. and McKay, H. D.: 1969, *Juvenile delinquency and urban areas: A study of rates of delinquency in relation to differential characteristics of local communities in American cities*, University of Chicago Press, Chicago.

Steffensmeier, D. and Allan, E.: 1996, Gender and crime: Toward a gendered theory of female offending, *Annual review of sociology* **22**, 459–487.

Sutherland, E. H.: 1947, *Principles of criminology*, University of Chicago Press, Chicago.

Sykes, G. M. and Matza, D.: 1957, Techniques of Neutralization - A Theory of Delinquency, *American sociological review* **22**, 664–670.

Van Wilsem, J., Wittebrood, K. and De Graaf, N. D.: 2006, Socioeconomic dynamics of neighborhoods and the risk of crime victimization: A multilevel study of improving, declining, and stable areas in the Netherlands, *Social Problems* **53**, 226–247.

Walker, J. L., Ehlers, E., Banerjee, I. and Dugundji, E. R.: 2011, Correcting for endogeneity in behavioral choice models with social influence variables, *Transportation Research Part A-Policy and Practice* **45**, 362–374.

Weerman, F. M.: 2011, Delinquent Peers in Context: A Longitudinal Network Analysis of Selection and Influence Effects, *Criminology* **49**, 253–286.

Young, J. T. N., Barnes, J. C., Meldrum, R. C. and Weerman, F. M.: 2011, Assessing and Explaining Misperceptions of Peer Delinquency, *Criminology* **49**, 599–630.

A The construction of an instrument for the crime rate

It is often difficult to find good instruments for the endogenous variables in models with unobserved heterogeneity. It is therefore very useful that Bayer et al. (2004) have developed a procedure for constructing an instrument in the context of such a model for neighborhood sorting. In this appendix we follow their suggestion and develop a method for constructing an instrument for the crime rate in the model we use here. Start by observing that, according to the model, the expected crime rate is:

$$E(C_j) = \left(\sum_{i \in j} \frac{e^{\beta X_i + \gamma Z_j + \delta E(C_j) + \xi_j}}{1 + e^{\beta X_i + \gamma Z_j + \delta E(C_j) + \xi_j}} \right) / B_j. \tag{5}$$

where the summation is over all individuals living in neighborhood j and B_j is the total number of these individuals. When the model is estimated, the coefficients β, γ and δ are known. The unobserved heterogeneity terms residuals ξ are the residuals from the second stage procedure and when they are substituted into (5) the equation holds as an identity: the observed crime rate is exactly replicated by the estimated model. The instrument is computed by deleting the unobserved heterogeneity terms ξ from (5) and computing the expected crime rate implied by the resulting equation. We denote this counterfactual crime rate as $IE(C_j)$:

$$IE(C_j) = \left(\sum_{i \in j} \frac{e^{\beta X_i + \gamma Z_j + \delta IE(C_j)}}{1 + e^{\beta X_i + \gamma Z_j + \delta IE(C_j)}} \right) / B_j. \tag{6}$$

Note that the $IE(C_j)$'s appear also on the right-hand side of the equation. They can therefore be interpreted as the crime rates that would be observed in a hypothetical world in which unobserved heterogeneity is absent.

The $IE(C_j)$'s are the desired instrument for the $E(C_j)$'s. The $IE(C_j)$'s are by construction uncorrelated wit the unobserved heterogeneity terms xi_j. Moreover, they are, also by construction, probably very strongly correlated with the $E(C_j)$'s.

One complication associated with the suggested procedure is that (6) uses the estimated coefficients of the model, which can only be obtained through the use of the instrument. Bayer et al. (2004) therefore propose an iterative procedure in which one starts with an informed guess of the instrument values, then computes the coefficient estimates and use them to recomputed the instrument, et cetera.