SOCIAL INTERACTIONS AND CRIME REVISITED: AN INVESTIGATION USING INDIVIDUAL OFFENDER DATA IN DUTCH NEIGHBORHOODS

Wim Bernasco, Thomas de Graaff, Jan Rouwendal, and Wouter Steenbeek*

Abstract—Using data on the age, sex, ethnicity, and criminal involvement of more than 14 million residents of all ages residing in approximately 4,000 Dutch neighborhoods, we test if an individual's criminal involvement is affected by the proportion of criminals living in his or her residential neighborhood. We develop a binomial discrete choice model for criminal involvement and estimate it on individual data. We control for both the endogeneity that may be related to unobserved neighborhood characteristics and for sorting behavior. We find significant social interaction effects, but our findings do not imply multiple equilibria or large multiplier effects.

I. Introduction

THE geographic variability of crime is a long-standing puzzle that was studied in the early nineteenth century by statisticians Quetelet (see Beirne, 1987) and Guerry (see Friendly, 2007). A seminal paper on the topic (Glaeser, Sacerdote, & Scheinkman, 1996) analyzes data on 658 cities in the United States and 70 precincts in New York City. The findings demonstrate that for a variety of crime types, the geographic variability in crime rates cannot be explained by economic, social, or legal differences between cities or precincts. The authors conclude that the remaining variability should be attributed to "social interactions," a term that encompasses a variety of different nonmarket mechanisms but is seldom explicitly defined (Manski, 2000). A common theme in the literature on social interactions is the proposition that the optimal choice of an individual depends on the choices of others, in particular others with whom the individual interacts directly or vicariously. If people interact predominantly with others who are geographically nearby, these interactions may provide an alternative explanation for the geographical variability of crime.[1]

This paper reconsiders the role of social interactions in crime using comprehensive and detailed data. One important advantage of our data is that they allow us to measure the potential sphere of influence of social interactions more precisely than has been possible in prior work on social interactions in terms of both spatial scale and reference group. With respect to spatial scale, we use Dutch neighborhoods that have an average population of 4,000 residents and an average surface of 10 square kilometers (approximately 4 square miles), whereas most previous studies applied their model to larger entities, such as cities and precincts. The detailed spatial scale is not only important for statistical reasons (e.g., it ensures less heterogeneity within and more heterogeneity between observations) but also from a substantive point of view. In order for the choices of individuals to be affected by those of others in their environment, they must be aware of these choices. Most individuals are much more likely to be aware of the behavior of residents in their own neighborhood than of the behavior of residents in remote parts of their city or region. In addition to measurement at a detailed spatial scale, we also consider the possibility that social interactions may be age specific. Because age segregation is endemic in all societies (Hagestad & Uhlenberg, 2006), and consequently most individuals are more aware of the behaviors of peers in their own age groups than of those who are much older or much younger, and because the criminological literature emphasizes the role of peers in criminal decision making, we also test age-specific social interactions.[2]

A second advantage of our data is that they include individual choices. Our data refer to individual neighborhood residents, which allows us to analyze offender rates (percentages of neighborhood residents suspected of being involved in crime in a given year) rather than crime rates. This allows us to separate crimes from offenders, that is, choices from agents. This contrasts with the study by Glaeser et al. (1996), which was based on crime rates—annual numbers of crimes per capita committed within the geographic boundaries of cities or precincts—forcing the authors to make assumptions on the number of crimes committed per criminal.

Another important advantage of our data is that they facilitate the inclusion of individual characteristics (sex, age, and ethnic origin) that are strongly correlated with criminal involvement. This allows us to take into account stylized facts about individual determinants of criminal conduct, such as the overrepresentation of males, adolescents, and nonnative residents among the criminal population. Our analysis includes the age (sixteen categories), sex (male versus female), ethnic origin (native versus foreign), neighborhood of residence (4,028 neighborhoods), and criminal record of the complete registered 2006 population aged 10 to 89 of the Netherlands. The sample totals 14.3 million

Received for publication April 9, 2014. Revision accepted for publication September 12, 2016. Editor: Bryan S. Graham.

* Bernasco and Steenbeek: Netherlands Institute for the Study of Crime and Law Enforcement; de Graaff and Rouwendal: Vrije Universiteit Amsterdam.

We thank the Dutch Police (KLPD) for providing access to crime data, the participants of the Vth Congress of the Spatial Econometric Association and the 52nd European Congress of the RSAI and two anonymous referees for comments that greatly improved the paper.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00656.

[1] See, for instance, Case and Katz (1991), Freeman, Grogger, and Sonstelie (1996), and Zenou (2003) and the references cited in these papers for further empirical evidence; Ballester, Zenou, and Calvó-Armengol (2010) for a theoretical analysis of delinquent networks in the tradition of Becker (1968), and Calvó-Armengol and Zenou (2004) for the presence of multiple equilibria.

[2] Alternatively, social interactions might include role model effects. For example, adolescents may copy the behavior of young adults. As this possibility is not highlighted in the literature and because including it would add another layer of complexity to the analyses, we leave it for future work to explore this possibility.

The Review of Economics and Statistics, October 2017, 99(4): 622–636

doi:10.1162/REST_a_00656

residents, of whom just over 21,300 were registered for being suspected by the police of having committed a crime in 2006.

As is well known, the identification of endogenous social interactions is complicated by the reflection problem (Manski, 1993), which "arises when a researcher observing the distribution of behavior in a population tries to infer whether the average behavior in some group influences the behavior of the individuals that comprise the group" (p. 532).

Several approaches to identify social interactions have been used in the literature. For example, Glaeser et al. (1996) develop a model with three types of agents—die-hard law breakers, die-hard law abiders, and those whose behavior depends on that of a close neighbor—and take the predictions of this model to their aggregate data. Bertrand, Luttmer, and Mullainathan (2000), who study welfare use, exploit differences between language groups and locations to identify network effects. Our empirical work is based on a binomial logit model of an individual's choice of whether to be a criminal. Explanatory variables include individual as well as neighborhood characteristics, one of which is the fraction of criminals residing in the neighborhood. Brock and Durlauf (2001) provide a set of conditions under which such endogenous social interactions are identified in binomial discrete choice models and can be separated from contextual effects. We employ the methodology of Berry, Levinsohn, and Pakes (1995), making use of a constructed instrumental variable inspired by the work of Bayer and Timmins (2007).

To anticipate our results, we find significant positive effects of social interaction, especially among young people, after controlling for the impact of unobserved neighborhood effects and eliminating possible effects of sorting. Further, we find that the strongest interaction effects are present for property crimes. However, our models do not suggest that multiple equilibria are relevant for the Dutch case. Nor do they imply large social multiplier effects of criminal behavior.

The remainder of this paper is structured as follows. The next section discusses the economic and criminological literature on crime and social interactions. Section III addresses the model and the estimation methodology. Section IV describes the data used. Section V presents the findings, after which the final section concludes and provides suggestions for future research.

II. Literature

In criminological research, it has long been observed that peer delinquency and individual delinquency are correlated—that those who break the law tend to associate with others who also break the law—although empirically the proposition is tested almost exclusively among juveniles, not adults. Two mechanisms have been hypothesized to underlie this correlation. The first mechanism is social learning (Akers, 1977; Sutherland, 1947), according to which criminal behavior is learned from delinquent peers. People are more likely to commit crime if peers also commit crime, and learning includes being taught the tangible techniques of committing crime but also learning cognitive techniques of neutralization to overcome moral concerns (Sykes & Matza, 1957). This mechanism is an example of contextual interaction with regard to learning specific skills, as well as endogenous interaction, because it implies behavioral interdependence (referring to the decomposition terms of behavioral similarity in groups by Manski, 1993).

The second mechanism is group selection. According to this argument, criminality itself is caused by other factors (such as weak social bonds or low self-control; see Gottfredson & Hirschi, 1990), and the propensity of an individual to be a criminal is not caused by the company of criminal friends. Instead, causality runs the other way: criminals tend to seek the company of other criminals. Because association is also based on geographical proximity (Festinger, Back, & Schachter, 1950), peer group or neighborhood selection induces behavioral similarity in criminality. This mechanism is an example of correlated effects, as it is not driven by behavioral dependence with respect to the decision to become a criminal.

The correlation between peer delinquency and individual delinquency is thus hypothesized to be affected by processes of selection and influence in social interactions between peers. In one of the first criminological 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 one's friends in the school network has a significant, although relatively small, effect on individual delinquent behavior. Patacchini and Zenou (2012) also study delinquency in peer networks and find a "conformism" effect of peers' delinquency for all crimes, but especially for petty crimes.

Social influence thus seems to be important to explain the correlation between peer delinquency and individual delinquency. Measuring the full extent of social networks to identify and estimate social interactions may, however, be unnecessarily demanding, because social interactions are likely to play a role not only in networks of strong ties but also in networks of weak ties. Social interactions may even work vicariously, including 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 (e.g., vandalism), and inferring the behavior.

In this paper, we focus on the endogenous interactions between neighborhood residents and test the hypothesis that, all other things being equal, an individual's decision to be a criminal positively depends on the proportion of neighborhood residents who are criminals. Thus, we expect that one's behavior is influenced by observing or learning about the behavior of other neighborhood residents. Relevant examples for the purposes of this paper are (a) see crime take place, (b) hear about crime taking place from offenders or victims in one's peer group, (c) see the results of crime, and (d) become a victim of crime. The neighborhood is an important context for studying the role of social interactions in crime. The focus on neighborhood as the presumed geographic unit of analysis where individuals interact is logical, given the wealth of published research on neighborhood effects (for an overview of outcomes unrelated to crime, see Sampson, Gannon-Rowley, and Morenoff, 2002).

We analyze property crime and violent crime separately as well as jointly in an overall measure that includes both types of crime. The rationale for separating the two is that if criminal social interactions exist, they may be crime type specific. This would mean that an individual's decision to become involved in property crime depends on the proportion of property offenders (but not on the proportion of offenders of violent crime) in his or her environment and that the probability of becoming involved in violent crime depends on the proportion of violent offenders (but not on the proportion of property offenders) in the area. Violent crime includes offenses like homicide, assault, vandalism, sexual assault, and robbery. Property crime includes offenses like burglary, larceny, theft, and arson. Because violent crime has a strong reciprocal nature (assault often takes place for reasons of revenge, and the perpetrators of assault are often identified while those of property crime often are not), we hypothesize that the social interaction effect for violent crime is larger than for property crime.

III. The Model

This section presents the model and the method of estimation. We use a binomial logit model for the choice whether to be a criminal. This choice is determined by personal characteristics as well as neighborhood characteristics. Idiosyncratic differences in individual choice behavior are captured by the conventional logit error term. We also address unobserved neighborhood effects by introducing elements of the approach pioneered by Berry et al. (1995) along the lines of Walker et al. (2011) in their model of social interactions in travel mode choice. Subsequently, we deal with the issue of identification of the social interaction effect, and we address the endogeneity of the social interaction effect. The final subsection discusses an implication of our model: the existence of multiple neighborhood offender rate equilibria.

A. *The Choice of Whether to Be a Criminal*

The model we use focuses on an individual's choice to become a criminal. This choice is conditional on the neighborhood in which the individual resides; we think of it as being based on a comparison of the utilities of being a criminal and being a law abider. We denote the difference between these two utilities for individual i living in neighborhood j as y_{ij}. It is important to note that it is only the difference between these two utilities that determines the decision to become a criminal. One may think of the utility of not being a criminal as the outside option in our model. The utility of this outside option does not need to be specified and is allowed to vary over the neighborhoods and individuals. It may depend, for instance, on labor market opportunities.

The variable y_{ij}, and, henceforth, the choice to become a criminal, depends on personal characteristics X_{ij} and neighborhood characteristics Z_j. A social interaction effect is present if the expected value of the offender rate in neighborhood j, C_j, has an impact on the probability that a particular individual i living in that neighborhood chooses to be a criminal.[3]

Let C_{ij} be a 0-1 variable that indicates whether individual i living in neighborhood j is a criminal. The probability that C_{ij} equals 1 is the probability that the variable y_{ij} is positive. We assume that y_{ij} is linear in the (observed) personal and neighborhood characteristics. Since we are imperfectly informed about these characteristics, we also introduce two random variables: ϵ_{ij} for unobserved personal characteristics and ξ_j for unobserved neighborhood characteristics. We assume that these variables (i.e., ξ_j and ϵ_{ij}) are independent of the personal characteristics X_{ij}. Moreover, we assume that ξ_j is independent of Z_j and, conditional on X_{ij}, ϵ_{ij} is independent of all neighborhood-specific attributes Z_j, $E(C_j)$, and ξ_j. The variable y_{ij} is thus specified as

$$y_{ij} = \alpha X_{ij} + \beta Z_j + \gamma E(C_j) + \xi_j + \epsilon_{ij}. \tag{1}$$

We assume that the variable ϵ_{ij} is logistically distributed, so that the probability that C_{ij} equals 1 is given by the logit expression

$$\Pr(C_{ij} = 1 | X_{ij}, Z_j, E(C_j)) = \frac{e^{\alpha X_{ij} + \beta Z_j + \gamma E(C_j) + \xi_j}}{1 + e^{\alpha X_{ij} + \beta Z_j + \gamma E(C_j) + \xi_j}}. \tag{2}$$

Note that $\Pr(C_{ij} = 1 | X_{ij}, Z_j, E(C_j))$ refers to an individual i, whereas $E(C_j)$ refers to the population of the neighborhood. $E(C_j)$ can be interpreted as the expected value that a randomly chosen individual living in j is a criminal.[4]

Without the social interaction and unobserved neighborhood effects ($\gamma = \xi_j = 0$), 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 an individual's probability to become a criminal but are unobserved by the analyst. The importance of such unobserved heterogeneity in discrete choice models is analyzed thoroughly by Berry (1994) and Berry et al. (1995) in their seminal study of the automobile market. Their approach is used in other fields as well. For instance, Walker et al. (2011) apply a model like equation (2), but without neighborhood variables Z_j, to study the effect of social interactions on travel mode choice.

[3] C_j is a neighborhood characteristic, but since it is the focus of interest of this paper, we set it apart from the other neighborhood characteristics Z_j.

[4] See equation (5). The value of $E(C_j)$ is conditional on the population living there and their characteristics, but for ease of notation, we have not made this explicit.

B. 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 δ_j,

$$\Pr(C_{ij} = 1|X_{ij}, \delta_j) = \frac{e^{\alpha X_{ij}+\delta_j}}{1 + e^{\alpha X_{ij}+\delta_j}}, \tag{3}$$

and this binary logit regression of C_{ij} on X_{ij} and a neighborhood dummy is estimated by maximum likelihood, based on the assumptions listed above. In the second stage, the alternative specific constants are analyzed further by writing them again as

$$\delta_j = \beta Z_j + \gamma E(C_j) + \xi_j. \tag{4}$$

The unobserved heterogeneity terms ξ_j are now the residuals of the linear regression equation. Manski (1993) studies identification of 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) show that the nonlinearity that occurs in a discrete choice model like equation (2) has identifying power. They develop a set of conditions under which all the remaining parameters are identified. These conditions apply to the model, equation (2), when the term referring to unobserved heterogeneity is absent.

Model (2) is identified if the parameters α and δ in equation (3) are identified and if the parameters β and γ in equation (4) are identified. Manski (1988) shows that the multinomial logit model is identified, so α and δ in equation (3) are not a problem. In Manski's (1993) linear model, C_{ij} is on the left-hand side of the linear equation of interest, whereas in equation (4), it is inside the estimated neighborhood-specific constant δ_j. This is the reason Manski's reflection problem does not occur in the present context. This implies that γ in equation (4) does not cause the same problem as in Manski (1993).

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 $E(C_j)$. $E(C_j)$ is therefore expected to be correlated with ξ_j. The reason is that a high value of ξ_j makes it more likely that any individual in the neighborhood is a criminal, which tends to increase $E(C_j)$. Hence the error term in equation (4) is not independent of the explanatory variables. In section IIIC, we propose a solution to this problem using an instrumental variable approach.

C. Endogeneity

As an instrument, we need an additional variable that is correlated with $E(C_j)$ but orthogonal to ξ_j. Since it is difficult to find such variables,[5] we will construct such an instrument on the basis of the structure of the model, adapting an idea that was developed originally by Bayer, McMillan, and Rueben (2004). The basic idea is that our model can be used to predict the offender rates that would be observed if there would not be unobserved heterogeneity—that is, if all ξ_j's would be equal to 0. These predicted offender rates are, by construction, independent of the ξ_j's, are likely to be strongly correlated with the actual, observed offender rates if the exogenous variables are salient, and make use of exogenous information that is not present in equation (4), namely, the personal characteristics X_{ij} of the inhabitants of the neighborhoods.[6] Moreover, these counterfactual offender rates clearly satisfy the exclusion restriction.

To see how the procedure works, start by observing that according to the model, the expected offender rate is

$$E(C_j) = \left(\sum_{i \in j} \frac{e^{\alpha X_{ij}+\beta Z_j+\gamma E(C_j)+\xi_j}}{1 + e^{\alpha X_{ij}+\beta Z_j+\gamma E(C_j)+\xi_j}} \right) \Big/ B_j, \tag{5}$$

where the summation is over all individuals living in neighborhood j and B_j is the total number of these individuals. It is easy to verify that in equation (5), there is a positive correlation between the unobserved neighborhood characteristics and the offender rate.

If we know the true values of the coefficients α, β, and γ, we would be able to compute counterfactual choice probabilities, denoted as $IE(C_j)$'s, for the situation in which unobserved neighborhood effects were absent, that is, for a situation in which all ξ_j's are equal to 0. The $IE(C_j)$'s are, by construction, uncorrelated with the ξ_j's, probably highly correlated with the $E(C_j)$'s, and satisfy the exclusion restriction. We thus use the exogenous characteristics X_{ij} and Z_j of individuals and neighborhoods to compute counterfactual choice probabilities for each individual that jointly predict counterfactual offender rates that are independent of the unobserved neighborhood characteristics.

The instrument is thus computed by removing the unobserved heterogeneity terms ξ_j from equation (5) and computing the expected offender rate implied by the resulting equation:[7]

[5] For instance, Walker et al. (2011) propose 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 socioeconomic characteristics. However, these approaches are easy to criticize. It is not difficult to imagine social interactions that cross the often somewhat arbitrary boundaries of neighborhoods, which would violate the exclusion restriction. It is also quite conceivable that the demographic composition of a neighborhood has a direct impact on the probability that some of its inhabitants become criminals.

[6] We report the first stage of one of the 2SLS regressions in the online appendix.

[7] Since we observe the whole population, we regard $IE(C_j)$ as a population quantity.

$$IE(C_j) = \left(\sum_{i \in j} \frac{e^{\alpha X_{ij} + \beta Z_j + \gamma IE(C_j)}}{1 + e^{\alpha X_{ij} + \beta Z_j + \gamma IE(C_j)}} \right) \Big/ B_j. \quad (6)$$

In the cases we consider, this equation always has at least one equilibrium. Although there can be multiple equilibria, the value of the instrument can always be defined unambiguously, as we discuss in online appendix A.

A complication associated with implementing the suggested procedure is that equation (6) uses the true (estimated) coefficients of the model, which can only be obtained through the use of the instrument. Bayer and Timmins (2007) therefore propose an iterative procedure in which one starts with an informed guess of the instrument values,[8] then computes the coefficient estimates, and use them to recompute the instrument until convergence is achieved (see also Bayer et al., 2004). In practice, this procedure works well. Convergence always occurred quickly, and simulation exercises confirm that the final values of the estimated coefficients are independent of the starting values in all cases considered. (See online appendix A for a detailed discussion.)

A possible concern with this procedure is that not all of the characteristics X_{ij} and Z_j are exogenous. It is conceivable that some of them are correlated with the unobserved neighborhood characteristics. We come back to this in section V.

D. Social Interaction and Multiple Equilibria

The implications of the presence of social interaction in our choice model at the neighborhood level can be investigated on the basis of equation (5). We can interpret the right-hand side of this equation as a mapping of $E(C_j)$ into itself. To focus on essentials, we assume a neighborhood populated by individuals who are identical[9] (apart from the idiosyncratic term in the logit model) and simplify equation (5) as

$$E(C_j) = \frac{e^{\phi_j + \gamma E(C_j)}}{1 + e^{\phi_j + \gamma E(C_j)}}. \quad (7)$$

In this equation, ϕ_j summarizes all other neighborhood and individual characteristics. It is not difficult to verify that

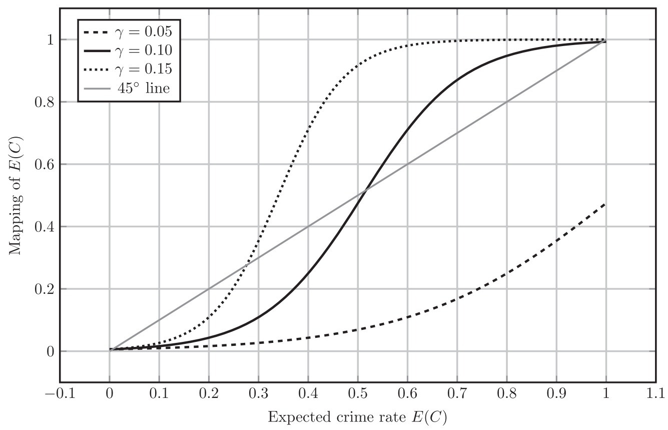
$$\frac{dE(C_j)}{d\phi_j} = \frac{1}{1 - \gamma E(C_j)(1 - E(C_j))} E(C_j)(1 - E(C_j)). \quad (8)$$

The first term on the right-hand side is a multiplier that equals 1 if there is no social interaction ($\gamma = 0$) and is larger than 1 whenever there is positive social interaction ($\gamma > 0$). As noted above, this multiplier could be responsible for spatial variation in offender rates much larger than one would expect on the basis of a model without endogenous social interactions.

[8] One can, for instance, use the OLS estimates of equation (4).

[9] In online appendix A, we discuss the (modest) changes that occur when the population is heterogeneous.

FIGURE 1.—POSSIBLE EQUILIBRIA FOR MULTIPLE FORMS OF CRIME BASED ON $\phi_j = -5.1$ AND VARIOUS γ VALUES



Brock and Durlauf (2001) provide an analysis of the equilibria in this model. They show that for positive values of γ, there may exist three equilibria. Figure 1 illustrates this situation for $\phi_j = -5.1$ (based on average estimation results we discuss below) and various values of γ. Equilibria occur where the curves cross the 45 degree line. For $\gamma = 0.1$ and $\gamma = 0.15$, there are indeed three equilibria. The high and stable equilibrium refers to a situation in which almost everybody is involved in criminal activities. In the other two equilibria, criminals are a minority, but the size of the minority differs significantly. The proportion of criminals equals either 1.5% or 27% (for $\gamma = 0.15$) or 52% (for $\gamma = 0.10$). The 1.5% equilibrium is stable, but the other two are not. The next section shows that the average offender rate in the Netherlands lies around 1.5%. However, the next section shows as well that there is substantial spatial variation in offender rates. This example therefore suggests that the model can be consistent with the presence of a different proportion of criminals in neighborhoods who are similar in all characteristics. Although the exact location of the equilibria depends on individual and neighborhood characteristics, our numerical experiments suggest that the offender rate at the stable equilibrium with the highest offender rate is unrealistically high.

IV. Data

Criminal behavior is notoriously difficult to measure. Because it is morally objectionable and legally sanctioned, many people are unwilling to confess their involvement in crime, to law enforcement as well as to researchers. Although quite a few surveys ask adolescent subjects to report their involvement in criminal conduct (a few examples include Elliott, Huizinga, & Ageton, 1985; Farrington et al., 1996; Wikström et al., 2012), crime self-report surveys are rare among adult populations (but see Morselli & Tremblay, 2004).

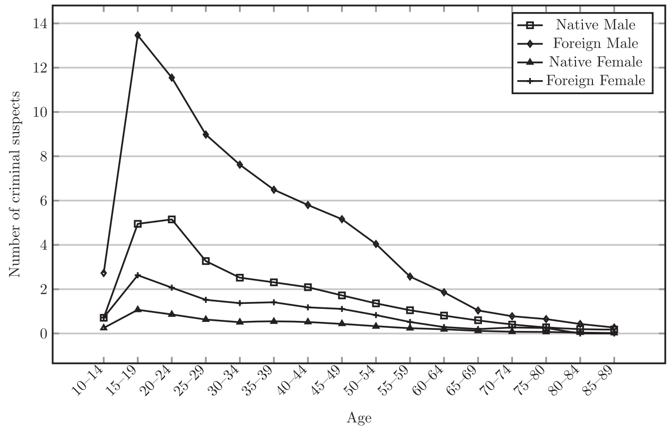
To measure criminal behavior, we therefore used anonymized national population data from the Dutch National

Police. The police information system from which the data were extracted contains data on all individuals who have been arrested by the Dutch police as criminal suspects in a particular year. It is estimated that more than 90% are subsequently either convicted in court or receive a fine or community service by the public prosecutor's office in lieu of prosecution (this often happens in the case of relatively minor crimes) (Blom et al., 2005). The data contain some personal characteristics (sex, age, country of birth, parents' countries of birth, postal code of residential address) and also 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 being a suspect of any crime(s) in the year 2006 as the dependent variable, as well as separate indicators for being suspected of violent and property crime. Violent crime includes homicide, assault, vandalism, sexual assault, and robbery; property crime includes burglary, larceny, theft, and arson. The two types do not exclude each other, so a single person can be suspected of both crime types within the same year. Although information about individual crime frequency is available, only the contrast between criminals (one or more crimes) and noncriminals (no crimes) was used in the analysis. This is unlikely to have any effects on outcomes because the overall variation in criminal involvement is almost completely captured by the distinction between criminals (1.3% of the population) and noncriminals (98.7% of the population).

Because the police information system is used for investigative purposes, it is updated continuously; 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 and 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 well-documented disadvantages of using police records to measure criminality. First, a substantial percentage of crimes never come 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, Lynch, & Nieuwbeerta, 2004). Second, in most jurisdictions, the police solve only approximately 20% 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 that the police use 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, but some of these people may be unjustly suspected and will not be convicted subsequently in court. Notwithstanding these limitations police records are the best available large-scale measures of criminality

Figure 2.—Number of Criminal Suspects in 2006 per 1,000 Individuals, by Age, Sex, and Ethnic Origin



available and have been used extensively in previous studies in the Netherlands and abroad.

To obtain a full population data set on criminal involvement in 2006 in the Netherlands, we used population data from Statistics Netherlands as of January 1, 2006, which cross-tabulates neighborhood of residence (4,028 neighborhoods) with age (twenty categories, each five years width), sex (male versus female), and ethnicity (native versus nonnative). As the police records contain these four variables as well, both sources can be combined to create a national data set containing approximately 16 million individuals. Because in the Netherlands only individuals of age 12 and older can be prosecuted, age categories 0 to 4 years and 5 to 9 years were removed from the analysis. Persons aged 10 or 11 are included because the population data are available only in five-year age categories. Because no individuals above age 89 were prosecuted in 2006, ages 90 and above were also removed from the analysis. The remaining data set contains 14,301,005 individuals aged 10 to 89 in 2006.

For this population, figure 2 displays the number of individuals who were suspected of criminal involvement during the year 2006, per 1,000 residents of the same sex, age category, and ethnic origin.[10] The figure confirms three stylized facts about criminality: the arrest rates of men are five times larger than those of women (Mears, Ploeger, & Warr, 1998; Steffensmeier & Allan, 1996), the arrest rates of residents with foreign origin are more than three times larger than those of native Dutch residents (Blokland et al., 2010), and the arrest rates of all groups peak during adolescence and early adulthood at ages 15 to 24 (Blokland, Nagin, and Nieuwbeerta, 2005). On average, 1.5% of the age 10 to 89 population became a crime suspect in 2006. For boys in the age category 15 to 24 years, the percentage is more than four times larger than the average.

The police records include the six-digit postal codes of the residential addresses of the individuals. Throughout the

[10] The data underlying figure 2 are included in table B.1 in the online appendix B.

FIGURE 3.—PERCENTAGE OF CRIMINAL SUSPECTS IN 2006 AMSTERDAM NEIGHBORHOODS

Source: HKS; maps Kadaster/Centraal Bureau voor de Statistiek (2013).

Netherlands, there are about 435,000 six-digit postal code areas. In nonrural areas they are roughly the size of a football field and contain approximately twenty residential properties and forty residents. Because they were created with pedestrian postal delivery services in mind, single codes are nearly always on the same street, apply to adjacent properties, and are not subdivided by physical barriers that impede pedestrian or car transportation. The focus of our investigation is the proportion of neighborhood residents involved in crime.[11] 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—a spatial aggregation of the six-digit postal code. 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" (p. 368). Many other studies in the Netherlands have used the four-digit postal code as a neighborhood delineation criterion (e.g., Bernasco & Kooistra, 2010; Nieuwbeerta et al., 2008; Wilsem, Wittebrood, & de Graaf, 2006).

Substantive arguments for neighborhood as a valid spatial reference group were given in section II. There are also several methodological arguments in favor of the neighborhood (instead of a larger or smaller spatial unit). First and foremost, larger areas such as cities ensconce within-city heterogeneity (and therefore between-neighborhood differences) in criminality. Second, smaller areas than neighborhoods, such as streets, result in very skewed crime distributions that are more difficult to model properly. Third, no or very few areal data are available at smaller spatial scales than neighborhoods. Because previous scholars have used cities as units of analysis, we bolster our argument for smaller areal units by presenting figure 3, which provides a view of neighborhoods in Amsterdam, the capital of the Netherlands. These figures show that the percentage of criminal suspects per municipality disguises large within-municipality differences. For example, whereas 2.2% of the population of Amsterdam was suspected of a crime in 2006 (1.8% and 1.2% for violent crime and property crime, respectively), the percentage of suspected criminals per neighborhood ranges from 0% to about 5% (0–4% and 0–3% for violent crime and property crime, respectively).

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 kilometers and an average population of 4,073 inhabitants. Similar to U.S. census tracts, the sizes of these "neighborhoods" depend on population

[11] When the peer group is located in 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 percentage of residents per square mile exhibiting a given behavior.

density. In urban areas where population densities are high, the surfaces of neighborhoods tend to be relatively small, whereas they are larger in rural areas where population densities are low.

To account for possible spurious findings, we control for several well-known correlates of crime in criminological research about neighborhood differences in crime. Classic and contemporary criminological studies have consistently found that high population turnover, ethnic heterogeneity, low socioeconomic status, and the presence of one-parent households correlate with higher offender rates (e.g., Bursik & Grasmick, 1993; Glaeser & Sacerdote, 1999; Sampson & Groves, 1989; Sampson, Raudenbush, & Earls, 1997). These neighborhood characteristics are hypothesized to affect crime in two distinct ways: by decreasing social cohesion and (expectations of) social control by impeding proper socialization of youth by parents and other neighborhood residents. We therefore merge our data set with two additional neighborhood data sets.[12] The first is the neighborhood data from the Statistics Netherlands from which we extract address density, percentage single-person households, average household size, number of shops, percentage owner-occupied housing, school density, and percentage single-parent households. The second is the Geomarketing data from WDM Netherlands, which is in itself composed of several (marketing) databases. This database gives us information about neighborhood mobility (in- and out-migration), average level of education, a measure for the average social class, and the number of double-income households within a neighborhood.

In addition, we aim to control for the effect of the presence and activities of law enforcement agencies. To that effect we use data from the Dutch Police Population Monitor of 2005, a biannual (cross-sectional) victimization survey spanning the entire Netherlands ($n = 52{,}560$). Several questions in this survey deal with the perception of the neighborhood by its residents. We aggregated the individual survey responses about the "perceived availability of the police" in their neighborhood to the four-digit postal code level. Perceived availability of the police is a combination of agreement ratings of five statements: "You don't see the police enough in this neighborhood," "They don't exit their vehicles often enough," "They are not approachable enough," "They do not have enough time for many things," and "They don't come quickly when you call them" (0 = agree, 1 = don't agree, don't disagree, 2 = disagree, with "don't know" recoded to answer category 1). The responses to these five questions were summed, resulting in a scale ranging from 0 to 10 (higher values indicating better availability). The mean of these individual-level summary values per four-digit postal code is the final availability of police per neighborhood variable ($n = 3{,}356$, mean = 4.8, sd = 1.5).

Finally, it is to be expected that the impact of social interaction varies with residential density. Denser neighborhoods presumably lead to more interaction among residents because meeting probabilities are larger.[13] To control for this effect, we incorporate an interaction effect between social interaction and residential density.

V. Results

This section presents our estimation results.[14] We start in section VA with the basic model of equation (3) and estimate it for the total population. Since it may be argued that social interactions are especially important among adolescents, who are also the most frequent offenders, we then estimate the model only for the group aged between 12 and 18. The results are discussed in section VB. Possible effects of sorting are probably less relevant for this group since most of them did not choose their own residential location. However, sorting may still be a concern when parents' characteristics are correlated with those of their children. As this is probably the case, we also present, in section VC, a version of the model in which indicators of neighborhood demographics have been replaced by averages over municipalities. The argument to support this specification is that choice of a residential neighborhood takes place within a larger (labor market) region that is taken as given by a household (see as well, e.g., Evans, Oates, & Schwab, 1992). This implies that the demographic characteristics of this larger area are also taken as given, whereas those of smaller geographical units are selected by the residential location choices of households. An additional advantage of the switch to averages of demographic characteristics of larger areas is that it removes the possible effects of differences in policing that are related to the neighborhood demographics (e.g., areas may have less or more intensive police surveillance) from the model. In this model, we also find significant endogenous social interaction effects for the total population as well as for the young.

A. The Basic Model

Table 1 presents the estimation results of equation (3) for the entire Dutch population, omitting the neighborhood-specific constants δ_j. The sociodemographic variables included are a sex indicator (0 for males, 1 for females), an ethnicity indicator (0 for native Dutch or born in a Western

[12] The actual number of resulting neighborhoods in our analyses is lower than the figure of 4,028 for two reasons. First, if the neighborhood is very sparsely populated (such as harbors and industrial areas), exogenous explanatory neighborhood variables such as age or household structure are not allowed to be reported because of privacy reasons and, the neighborhood drops out of our analysis. Second, if there is no criminal variation within a neighborhood (i.e., no criminals live in the neighborhood), then the neighborhood drops out of our analysis as well. The former does not lead to significant selection problems; the latter, however, might pose a potential problem, which we deal with in section V.

[13] See, for instance, Glaeser and Sacerdote (1999) for a discussion.

[14] The neighborhood-specific constants have been estimated by maximum likelihood with the Stata software package. Our second-stage models have been estimated by two-stage least squares using the R software package (R Core Team, 2014), most notably using the `AER` package (Kleiber & Zeileis, 2008).

TABLE 1.—CHOICE MODELS (LOG-ODDS OF BEING SUSPECT OF CRIME IN 2006)

	All Crime		Violent Crime		Property Crime	
Parameter	Estimation	SE	Estimation	SE	Estimation	SE
Female	−1.637	0.0059	−1.885	0.0069	−1.509	0.0085
Nonnative	0.783	0.0059	0.743	0.0063	0.957	0.0081
Age	0.070	0.0019	0.109	0.0021	−0.110	0.0030
Age^2	−0.026	0.0002	−0.030	0.0003	−0.034	0.0004
Number of observations	14,191,721		14,189,082		13,966,926	
Number of parameters	4 + 3,610 constants (δ)		4 + 3,602 constants (δ)		4 + 3,209 constants (δ)	
Log likelihood	−961, 220.0		−855, 499.9		−491, 169.2	

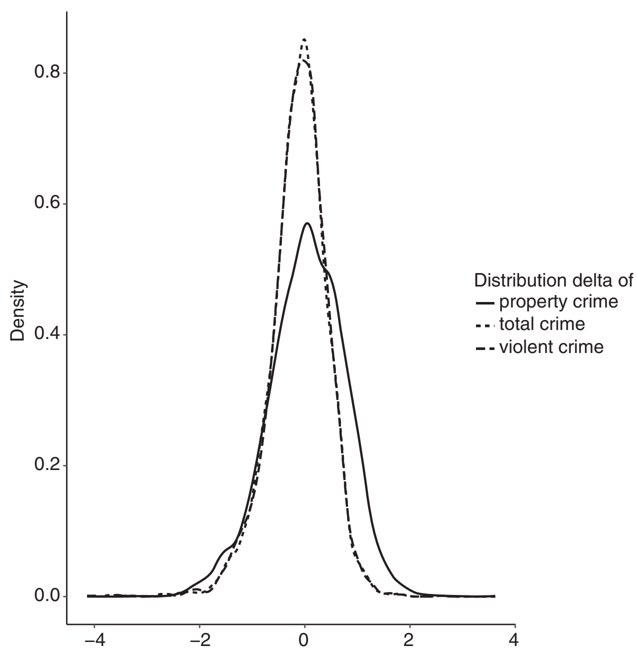
country, 1 for people or one of their parents born in a non-Western country), age (measured categorically as $10-14 = -1, 15-19 = 0, 20-24 = 1, \ldots, 85-89 = 14$—centered on the peak of the age-crime curve), and age squared. Of the 4,028 neighborhoods, there were 422 in which not a single resident offended in 2006, making it impossible to estimate a neighborhood-specific constant term for the general model. Violent offenders were absent in 430 neighborhoods, and property offenders were absent in 833 neighborhoods. Note that the number of observations decreases much more slowly than the number of neighborhoods, indicating that typically the sparsely populated neighborhoods drop out of the analysis. However, there are some nonsparse neighborhoods that show zero crime rates as well. To investigate the impact of this selection, section VD deals with a median regression with the nonsparse neighborhoods included as a sensitivity analysis.

The estimation results confirm the descriptive statistics visualized in figure 2. Males and nonnatives are much more likely to become involved in crime than females and native Dutch residents, and crime involvement quickly increases with age during adolescence and then gradually decreases. While the estimated parameters for violent crime are similar to the estimates for general crime, the property crime estimates indicate that the age-crime curve for property crime peaks at younger ages.

A higher value of the neighborhood-specific constant δ_j means that, conditional on sex, age, and ethnic origin, neighborhood residents are more likely to be involved in crime. The first-stage model imposes a structure on the effects of individual characteristics, but it is silent about the mechanisms underlying between-neighborhood variation: these have to be sorted out in the second stage. The kernel density estimates of the shape of the δ_j distributions are presented in figure 4. All three density functions are single peaked and almost symmetric. The kernel density function of violent crime is similar to that of crime in general, whereas the kernel density function of property crime clearly has a higher mean and standard deviation.

Table 2 presents the results of the second stage.[15] Of all three definitions of crime (all crime, violent crime, and

FIGURE 4.—KERNEL DENSITY ESTIMATES OF THE DISTRIBUTION OF δ



property crime), we first present the OLS result and then the 2SLS results.

For instruments of *Offender rate* and *Offender rate* × *density*, we used the counterfactual offender rate ($IE(C_j)$) as defined in equation (6), and the multiplicative interaction effect of the counterfactual offender rate and the address density ($IE(C_j)$ × *density*). Table B.2 in the online appendix reports the results of the first-stage regression on our social interaction variable (*Offender rate*) and interaction effect (*Offender rate* × *density*). As these tables clearly show, the instruments we use are relevant. The counterfactual offender rate is significant and shows a positive correlation with *Offender rate*, as expected. The counterfactual offender rate times the address density is significant as well and shows a positive relation with the interaction term: *Offender rate* × *density*. By construction, the counterfactual offender is uncorrelated with the error term ξ. We note that the way we compute our instruments implies that they are functions of the estimated parameters. This should in principle be taken

[15] The number of neighborhoods differ because specific neighborhood variables are missing due to privacy reasons and neighborhoods drop out because they contain no criminals. The former occurs in neighborhoods with a very small number of inhabitants, and the latter could occur as well in more populated neighborhoods. Section VD addresses the second cause of missing neighborhoods.

TABLE 2.—OLS AND 2SLS REGRESSIONS ON NEIGHBORHOOD-SPECIFIC CONSTANTS (δ)

		Neighborhood Demographics				Municipal Demographics			
		Total Sample		Between 12 and 18 years		Total sample		Between 12 and 18 years	
		1 OLS	2 2SLS	3 OLS	4 2SLS	5 OLS	6 2SLS	7 OLS	8 2SLS
All crime	Offender rate	0.603***	0.011	0.392***	0.056*	0.583***	0.047	0.381***	0.212***
		(0.007)	(0.039)	(0.001)	(0.029)	(0.007)	(0.034)	(0.005)	(0.014)
	Offender rate × address density ÷ 100	−0.401***	−0.014	−0.286***	−0.131***	−0.488***	−0.143***	−0.321***	−0.261***
		(0.017)	(0.049)	(0.019)	(0.050)	(0.018)	(0.052)	(0.019)	(0.039)
	Address density ÷ 100	0.583***	−0.020	0.858***	0.475**	0.865***	0.526	0.976***	0.994***
		(0.095)	(0.125)	(0.083)	(0.199)	(0.054)	(0.131)	(0.081)	(0.150)
	One-parent households	−2.452***	3.841***	−0.882***	3.565***	−2.414***	3.850***	−2.073***	0.627***
		(0.196)	(0.522)	(0.280)	(0.501)	(0.388)	(0.776)	(0.256)	(0.683)
	Observations	3,210	3,210	2,609	2,609	3,210	3,210	2,609	2,609
	Adjusted R^2	0.814	0.417	0.783	0.372	0.787	0.324	0.774	0.641
Violent crime	Offender rate	0.696***	0.043	0.478***	0.061	0.676***	0.004	0.469***	0.269***
		(0.008)	(0.044)	(0.006)	(0.074)	(0.008)	(0.046)	(0.006)	(0.021)
	Offender rate × address density ÷ 100	−0.438***	−0.008	−0.280***	−0.170**	−0.537***	−0.133**	−0.320***	−0.363***
		(0.020)	(0.059)	(0.024)	(0.074)	(0.021)	(0.067)	(0.024)	(0.061)
	Address density ÷ 100	0.520***	−0.046	0.607***	0.460**	0.796***	0.505***	0.709***	1.066***
		(0.052)	(0.132)	(0.081)	(0.221)	(0.053)	(0.147)	(0.080)	(0.176)
	One-parent households	−2.371***	2.983***	−0.701**	2.689***	−2.302	3.543***	−1.770***	−0.041
		(0.191)	(0.479)	(0.230)	(0.592)	(0.386)	(0.825)	(0.532)	(0.698)
	Observations	3,207	3,207	2,569	2,569	3,207	3,207	2,569	2,659
	Adjusted R^2	0.816	0.420	0.785	0.303	0.787	0.214	0.777	0.625
Property crime	Offender rate	1.093***	−0.242**	0.671***	0.111**	1.108***	0.224***	0.648***	0.344***
		(0.018)	(0.117)	(0.012)	(0.052)	(0.016)	(0.069)	(0.011)	(0.0285)
	Offender rate × address density ÷ 100	−0.768***	−0.034	−0.530***	−0.222**	−0.933***	−0.589***	−0.553***	−0.397***
		(0.039)	(0.129)	(0.038)	(0.093)	(0.004)	(0.102)	(0.038)	(0.074)
	Address density ÷ 100	0.636***	0.078	1.000***	0.586***	0.958	1.031***	1.019***	0.934***
		(0.070)	(0.182)	(0.105)	(0.224)	(0.071)	(0.150)	(0.010)	(0.174)
	One-parent households	−1.005***	8.406***	−0.936	4.051***	−1.861	4.963***	−2.010	2.243***
		(0.313)	(0.952)	(0.404)	(0.110)	(0.599)	(1.025)	(0.733)	(0.924)
	Observations	2,955	2,955	2,248	2,248	2,955	2,955	2,248	2,248
	Adjusted R^2	0.780	0.307	0.719	0.375	0.761	0.448	0.717	0.593

Standard errors in parentheses and controlled for neighborhood or municipality average variables as reported in table B.3 in the online appendix. Significance at $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

into account in the computation of the standard errors. However, since no methods for doing so appear to be available in the literature, we had to leave this issue unaddressed.

Columns 1 and 2 in table 2 present the most important results from our second-stage regression on neighborhood constants for our total sample and covariates measured at the neighborhood level. Table B.3 in the online appendix gives the full results of this regression, including the coefficients for the demographic control variables. The main conclusion is that the hypothesized social interaction effect (i.e., the effect of the neighborhood percentage of other residents involved in crime) that appears to be strong and highly significant in the OLS regressions, disappears completely in the 2SLS results. The estimates of most other variables in the 2SLS estimation change little with one exception: the coefficient for single-parent households was (unexpectedly) significantly negative in the OLS regression but becomes positive, larger in absolute value, and strongly significant in the 2SLS regressions. A large and highly significant coefficient on the share of single-parent households is in line with earlier analyses (notably Glaeser & Sacerdote, 1999).

The estimated coefficients of the other variables are in line with studies on neighborhood-level correlates on crime and delinquency, which generally show that indicators of social and economic disadvantages (low education, low income, high neighborhood mobility, and high proportions of single-person and single-parent households) are associated with more crime.[16]

B. *Only the Young*

In columns 3 and 4 in table 2, we report estimation results for a sample of persons between 12 and 18 years. The results of the first-stage (binomial logit) estimates are available from the authors on request. Here, we discuss only second-stage results. As we have seen, most crimes are committed by young people, which makes it interesting to give separate attention to this group. It may also be conjectured that social interactions are more important for these persons.[17] Finally, a focus on the young may to some extent alleviate potential concerns about sorting effects, as they did not choose their

[16] Our results are robust to the specification used. Only if we omit single-parent households do we get higher social interaction effects (up to $\gamma = 0.18-0.20$ for all and violent crime, respectively, and convergence problems for property crime). Using different instruments, in particular the spatial lags of surrounding neighborhoods, does, however, significantly increase the social interaction effect (with γ ranging from 0.54 to almost 1).

[17] There is evidence of strong age segregation in social contacts, see, for instance, Hagestad and Uhlenberg (2006).

residential neighborhood themselves. (See section VC for even further consideration of sorting effects.)

Columns 3 and 4 confirm the general conclusion from columns 1 and 2 in that an apparently strong social interaction effect decreases substantially once we take into account the possible impact of unobserved neighborhood characteristics. However, the *Offender rate* variable now remains (marginally) significant for all crime and for property crime. Moreover, we now find a significantly positive impact of density (addresses per hectare) and a significantly negative impact of the interaction between the offender rate and density for all types of crime. We therefore conclude that for the young, social interaction and an urban (high population density) environment are significant determinants of criminality.

C. *Municipal Demographics*

There are two other concerns with the estimates presented thus far that we investigate further in this section. We noted already that sorting may disturb our results. Offender rates may affect the composition of the population in neighborhoods, for instance, because high-income households avoid locating in areas with a high share of criminals and become underrepresented there, while other household types, for instance, those that experience more tight financial constraints, become overrepresented. If this happens, the share of criminals in the neighborhood becomes correlated with the demographic composition of that neighborhood, and this complicates the measurement of the determinants of criminality. It may even be argued that this effect is present if we consider only the young people who did not themselves choose their residential neighborhood, since their characteristics are correlated with those of their parents through nature and nurture. To address the sorting issue, Evans et al. (1992), who use binomial choice models to study peer effects in teenage behavior (pregnancy and school dropout), propose using averages over larger geographical areas. They argue that sorting refers mainly to the choice of a neighborhood within a metropolitan area, whereas households take the general demographic characteristics of this larger area as given. We follow their line of reasoning here and replace the demographic characteristics of the neighborhood by the average values of these variables in the municipality.[18]

This reformulation of the model addresses a second concern. Bertrand et al. (2000) motivate the use of area-wide averages of welfare use as a means to avoid possible bias associated with unobserved characteristics that individuals in a particular neighborhood have in common with others who belong to the same group. In the context of this paper, a potentially important effect is the attention given by the police to a particular neighborhood, which may well be correlated with the demographic characteristics of the neighborhoods. If the police have a given capacity per larger geographical unit, our alternative specification should be expected to be more robust against this possible source of bias as well.

The results of the alternative specification for the total population and for the young only are presented in table 2 in columns 5 and 6, and 7 and 8, respectively. For the young and for property crime, we now find highly significant endogenous social interaction effects. Although in the 2SLS regressions, these effects are smaller than indicated by simple OLS, they remain substantial. There is now also in all cases but one a highly significant positive effect of address density, while the interaction effect is negative. We conclude from these results that sorting as well as possible differences in the attention the police pay to various neighborhoods had an important impact on the earlier results. We regard the specification with neighborhood demographics as the preferred one.

D. *Sensitivity Analysis*

Median regression. Some neighborhoods drop out of our sample because there is no within-neighborhood crime variation; that is, no criminals live in those neighborhoods. For all crime and the total sample, only 2.4% of the relevant neighborhoods drop out of the sample, but for only the young and property crime, this percentage increases to 20.3%. To assess whether this makes a large impact, we therefore impute the value of -3.5 for all the missing δ's (this is slightly less than the minimum of all δ's we observe) with an accompanying standard error of 1, which is typically the standard error of neighborhoods with a very small number of criminals. We then compute the associated instruments as the expected offender rate implied by equation (6) and our estimates of α and β.

In our second stage, we estimate a median regression. To do so, we use the control function approach in quantile regression models as advocated by Lee (2007). The results can be found in table B.4 in online appendix B and are robust with respect to the value of the imputed δ's, the value of standard error of the δ's and the exact nonparametric form of the reduced-form residuals in the second stage of the analysis. In general, the median regression results for the offender rate are slightly more positive and significant (most notably for property crime and municipality average variables), but concur with the estimates in table 2 and confirm our results.

Alternative instruments. The instrument for the social interaction effect that we have proposed follows seamlessly the logic of our model and works well in practice. As we pointed out, the main issue is that this instrument brings in additional exogenous information: the personal characteristics of the inhabitants of various neighborhoods (compare the motivation of a similar instrument in Bayer & Timmins, 2007). However, it is possible to exploit this in other ways

[18] The variables concerned are: density one-parent households, average persons per household, average educational attainment, average social class, percentage double-income households, percentage in- and out-migration, percentage owner-occupied housing, and average on-street police presence.

than we have done. One possibility is to compute the counterfactual offender rates without using the social interaction effect (by setting $\hat{\gamma}$ equal to 0). This would remove the possibility of multiple equilibria, although this comes possibly at the cost of a loss in correlation between the instrument and the observed offender rates. Moreover, it addresses a concern that some may have with the computation of our instrument: that it could be weak if it uses only the exogenous information incorporated in X_{ij} and Z_j without incorporating the impact of the social interaction effect. We have therefore computed the instrument alternatively with the social interaction coefficient set equal to 0, keeping everything else in the procedure unchanged. The results of estimating the models of table 2 with this alternative instrument are reported in table B.5 in the online appendix. In almost all cases, there are only minor changes in the estimated coefficients. The single exception is model 2 for property crime, where the estimated coefficients for the field effect and its interaction with population density become much larger in absolute value (although negative).

One disadvantage of this approach is that the use of an iteration process for the counterfactual offender rates remains. We therefore computed an instrument where we set both $\hat{\beta}$ and $\hat{\gamma}$ at 0 when computing the instrument. These results can be seen in table B.5 in the online appendix. Again, in all cases, we see only minor changes in the coefficients when compared with tables 2 and B.5. Note that although we circumvent the iteration process, we still need imputed coefficients $\hat{\alpha}$ for the computation of the instrument. Although table 1 clearly shows that the coefficients $\hat{\alpha}$ are precisely estimated, one might still object to the use of any imputed coefficients in the calculation of our instrument.

Therefore, we go one step further and make no attempt to compute counterfactual offender rates, but use the personal characteristics of the inhabitants of the neighborhoods as instruments. This would imply the use of standard IV procedures. Using, for example, the age of the inhabitants of the neighborhood as an instrument yields qualitatively similar results as reported in tables 2 and B.5, although the standard errors are unsurprisingly larger due to the loss of correlation between the instrument and the observed offender rates.[19]

E. *Implications*

In this section we present the implications of our coefficient estimates for the equilibria. We start with investigating the implications of our model for the equilibrium offender rates. To do this we construct three diagrams similar to figure 1, but based on our estimation results and computed for the actual population, which is (of course) heterogeneous. Because of heterogenenous populations, these diagrams differ across neighborhoods. These figures specifically refer to the largest neighborhood (Molenvliet) of the small town of Woerden, whose characteristics are considered by marketers as average for the Netherlands. For a total overview of the equilibria for all neighborhoods for the youth and neighborhood-average exogeneous variables, we refer to figure A.2 in online appendix A.

FIGURE 5.—POSSIBLE EQUILIBRIA FOR MULTIPLE FORMS OF CRIME BASED ON PARAMETER ESTIMATES FOR THE YOUTH AND NEIGHBORHOOD-AVERAGE EXOGENEOUS VARIABLES FOR THE REPRESENTATIVE NEIGHBORHOOD OF MOLENVLIET IN THE CITY OF WOERDEN

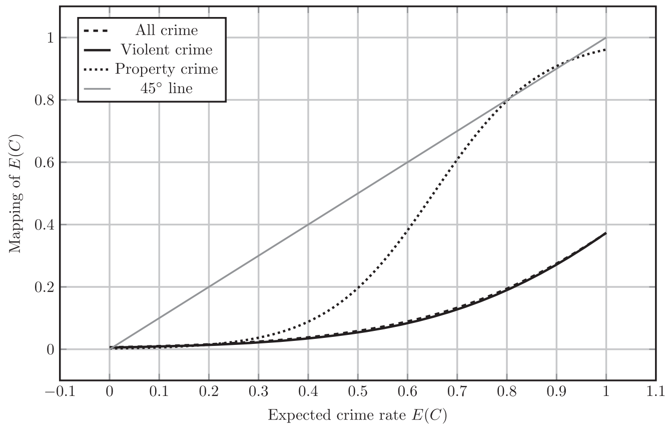


Figure 5 is based on the estimates listed in column 4 of table 2. It illustrates that multiple equilibria for property crime are possible in the model that refers only to the youth. In all three cases shown, there is an equilibrium close to the 1% or 2% crime rates that most of the neighborhoods exhibit; this particular neighborhood displays a 2.4% general crime rate, a 1.7% violent crime rate, and a 1.2% property crime rate. Here, the low equilibria and observed crime rates do not differ by more than 2.5% for more than 95% of the neighborhoods. For property crime, there is a second stable equilibrium at a very high level—close to 92%—of crime. The attraction basins of these two stable equilibria are separated by the third, unstable, equilibrium that is located above the maximum share of criminals in our data. These models therefore do not suggest that neighborhoods may switch from the low to the high equilibrium.

Figures 6 and 7 show the equilibria for our total sample and for youth only, respectively, implied by the models that use municipality average covariates (columns 6 and 8 in table 2). Now the attraction basins of the low-share equilibria are much smaller. Indeed, figure A.2 in online appendix A and analogous figures for other types of crime[20] show that unstable equilibria that indicate the upper bound of this basin are within the range of a few neighborhoods only, mostly with respect to property crime.[21] Although this may be interpreted as suggesting that some Dutch neighborhoods may run the

[19] In the case of total crime for the whole sample with neighborhood-specific demographics, the estimated offender rates in this case yield 0.086 with a standard error of 0.048. Full results are available on request.

[20] Available on request.

[21] Not surprisingly, the empirical distribution of neighborhood offender rates is highly skewed to the right. Ninety percent of all neighborhoods for all crime to more than 95% for property crime display offender rates below 5%. For all three types of crime, not more than four neighborhoods display offender rates above 10%.

Figure 6.—Possible Equilibria for Multiple Forms of Crime Based on Parameter Estimates for the Total Sample and Municipality-Average Exogeneous Variables for the Representative Neighborhood of Molenvliet in the City of Woerden

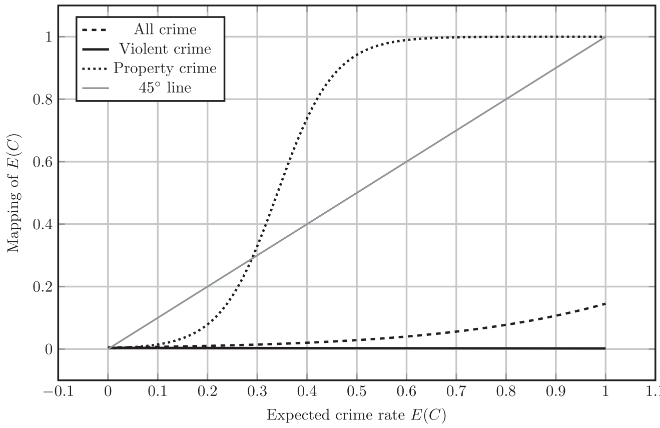
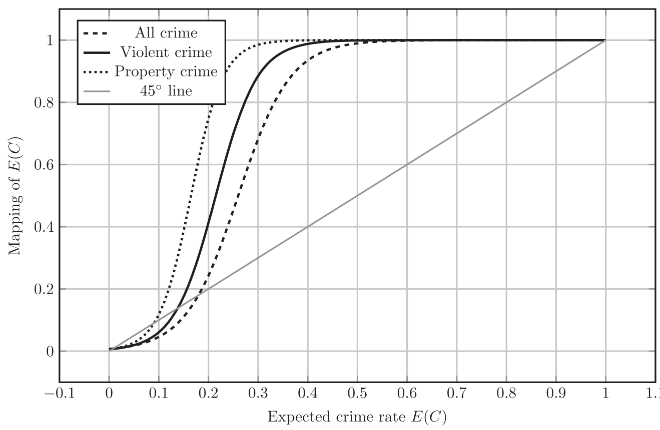


Figure 7.—Possible Equilibria for Multiple Forms of Crime Based on Parameter Estimates for the Youth and Municipality-Average Exogeneous Variables for the Representative Neighborhood of Molenvliet in the City of Woerden



risk of moving toward a much higher level of criminality, we do not observe such high offender rates in the data for any of the Dutch neighborhoods. We interpret this as indicating that (at least in the present circumstances) the lower-level equilibrium is the only relevant one.

The social multiplier of criminal behavior implied by our model was derived in equation (8). It gives the additional impact of a change in an exogenous variable on the equilibrium offender rate that is caused by the social interaction effect. Since the expression $E(C_j)(1-E(C_j))$ reaches a maximum value of 0.25 for $E(C_j) = 0.5$, whereas the offender rates we observe are in general much lower and our estimates of the field effect γ are well below 1, we should not expect large values of this multiplier. Computations reveal that for the model referring to youth only, with municipality averages of the demographic variables (column 8 of table 2) this multiplier reaches values in the range 1.018 to 1.031 (for neighborhoods with high offender rates, $E(C_j) = 0.1$), which is indeed small.

VI. Discussion

Social interactions may be caused by a variety of mechanisms, including contacts at school and work or through social and other media. In line with previous work on social interactions and with the comprehensive literature on neighborhoods and crime (see Sampson et al., 2002), our study focused on social interactions through coresidence and locational spillovers.

On the basis of previous literature, we hypothesized that positive social interactions play a role in crime—in particular, in violent crime. Using data on individual choices and taking unobserved neighborhood characteristics and sorting into account, our analysis of offender rates confirms this expectation. They indicate a positive and significant endogenous social interaction effect for crime overall, as well as for violent crime and property crime considered separately. However, our models do not suggest that multiple equilibria are relevant in the present Dutch circumstances. Nor do they imply important social multiplier effects of criminal behavior. This contrasts with some earlier research on social interactions and crime (Glaeser et al., 1996). We suggest that prior research may have overestimated endogenous social interaction effects by lack of individual data at a detailed spatial scale. In particular, we demonstrated that there exist huge individual differences in crime involvement by sex, age, and ethnic background, which have hardly been accounted for in prior research on social endogenous interactions and crime.

Nevertheless, in line with much of the prior literature, we did find positive social interaction effects in all three models once endogeneity and sorting effects were taken into account. The result for violent crime is easiest to interpret. Apart from violent crime's reciprocal nature, another interpretation might be that social interactions apply to violent crime because violent crimes are overt predatory contact crimes that presume an interaction between the offenders and their victims. In neighborhoods where individuals live among others who are prone to violence, the risk of violent victimization is relatively high and might be lowered by gaining a reputation of toughness. Thus, being a violent offender may deter violent predators and thereby prevent future violent victimization (Dur & van der Weele, 2013; Fagan & Meares, 2008; Silverman, 2004). The larger coefficient for the social interaction effect related to property crime is perhaps a bit more difficult to interpret. Although the large majority of property crimes (e.g., larceny, burglary) are covert crimes that are perpetrated without any contact between the perpetrator and the victim and without the victim being able to identify the perpetrator (and often also vice versa), it also covers crimes that have a social component. For instance, shoplifting may give the delinquent a reputation in his (or perhaps her) peer group that evokes similar behavior by others.

As we have pointed out already in this paper, there are some advantages as well as disadvantages of using police-recorded data on criminal behavior. First, some crimes are not reported to the police (more than 50% according to Goudriaan et al., 2004), and, second, the police solve only 20% of the recorded crimes (Dodd et al., 2004). Police data thus suffer from both type 1 and type 2 errors (although, as we have argued, the likelihood of a false positive is quite small). In future research, such misclassification issues might be dealt with. Lewbel (2000) has shown that binary discrete-choice models with misclassification are nonparametrically identified, and Hausman, Abrevaya, and Scott-Morton (1998) provide techniques for estimating this model.

A second issue that could be addressed further in subsequent research is that we found the share of single-parent households to be an extremely important variable. We were, however, unable to test the hypothesis suggested by this finding: that criminals often belong to such households since our data inform us only about the share of one-parent households per neighborhood. With better data, this issue could be addressed.

In a critique of the empirical literature on social interactions, Manski (2000) claims that the analysis would benefit from the performance of well-designed experiments in controlled environments and from careful elicitation of persons' subjective perceptions of the interactions in which they participate. Falk, Fischbacher, and Gächter (2010) adopt the first suggestion and demonstrate social interactions in an experiment on behavior in a public goods game. However, ethical considerations and Institutional Research Board regulations prohibit experimental studies of criminal behavior of the type and severity that we study. Therefore, in this paper, we chose the second-best alternative and estimated a structural discrete model using state-of-the-art techniques to tease out social interactions with an exceptionally rich and comprehensive data set. Our results confirm the presence of social interaction effects, but according to our estimates, they are not strong enough to make the presence of multiple equilibria in Dutch neighborhoods likely. The implied social multiplier is small, almost negligible. The bottom line of our findings therefore is that in our data, the variation in exogenous determinants of crime, like the demographic composition of the area and personal characteristics, appears to be more relevant for the explanation of geographical differences in offender rates than social interactions.

REFERENCES

Akers, Ronald L., *Deviant Behavior: A Social Learning Approach* (Belmont, CA: Wadsworth, 1977).

Ballester, Coralio, Yves Zenou, and Antoni Calvó-Armengol, "Delinquent Networks," *Journal of the European Economic Association* 8 (2010), 34–61.

Bayer, Patrick, Robert McMillan, and Kim Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market," NBER working paper w10865 (2004).

Bayer, Patrick, and Christopher Timmins, "Estimating Equilibrium Models of Sorting across Locations," *Economic Journal* 117 (2007), 353–374.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76 (1968), 169–217.

Beirne, Piers, "Adolphe Quetelet and the Origins of Positivist Criminology," *American Journal of Sociology* 92 (1987), 1140–1169.

Bernasco, Wim, and Thessa Kooistra, "Effects of Residential History on Commercial Robbers' Crime Location Choices," *European Journal of Criminology* 7 (2010), 251–265.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics* 25 (1994), 242–262.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica: Journal of the Econometric Society* 63 (1995), 841–890.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics* 115 (2000), 1019–1055.

Blokland, Arjan A. J., Kim Grimbergen, Paul Nieuwbeerta, and Wim Bernasco, "Criminaliteit en Etniciteit. Criminele carrières van Autochtone en Allochtone Jongeren uit het Geboortecohort 1984 [Crime and ethnicity: Criminal careers of native and nonnative youth of birth cohort 1984]," *Tijdschrift voor Criminologie* 52 (2010), 122–152.

Blokland, Arjan A. J., Daniel Nagin, and Paul Nieuwbeerta, "Life Span Offending Trajectories of a Dutch Conviction Cohort," *Criminology* 43 (2005), 919–954.

Blom, Martine, Ko Oudhof, Rob V. Bijl, and Bart F. M. Bakker, *Verdacht van Criminaliteit: Allochtonen en Autochtonen Nader Bekeken* [Suspected of crime: Nonnatives and natives examined more closely] (Heerlen/Den Haag: CBS/WODC, 2005).

Brock, William A., and Steven N. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies* 68 (2001), 235–260.

Bursik Jr., Robert J., and Harold G. Grasmick, *Neighborhoods and Crime: The Dimensions of Effective Community Control* (New York: Lexington Books, 1993).

Calvo-Armengol, Antoni, and Yves Zenou, "Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior," *International Economic Review* 45 (2004), 20.

Case, Anne C., and Lawrence F. Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth," NBER working paper 3705 (1991).

Dodd, Tricia, Sian Nichlas, David Povey, and Alison Walker, "Crime in England and Wales 2003/2004," *Home Office Statistical Bulletin* 10/04 (2004).

Dur, Robert, and Joel van der Weele, "Status-Seeking in Criminal Subcultures and the Double Dividend of Zero-Tolerance," *Journal of Public Economic Theory* 15 (2013), 77–93.

Elliott, Delbert S., David Huizinga, and Suzanne S. Ageton, *Explaining Delinquency and Drug Use* (Beverly Hills, CA: Sage, 1985).

Evans, William N., Wallace E. Oates, and Robert M. Schwab, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy* 100 (1992), 966–991.

Fagan, Jeffrey, and Tracey L. Meares, "Punishment, Deterrence and Social Control: The Paradox of Punishment in Minority Communities," *Ohio State Journal of Criminal Law* 6 (2008), 173–229.

Falk, Armin, Urs Fischbacher, and Simon Gächter, "Living in Two Neighborhoods: Social Interaction Effects in the Laboratory," *Economic Inquiry* 51 (2010), 563–578.

Farrington, David P., Rolf Loeber, Magda Stouthamer-Loeber, Welmoet B. Kammen, and Laura Schmidt, "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 (1996), 493–517.

Festinger, Leon, Kurt W. Back, and Stanley Schachter, *Social Pressures in Informal Groups: A Study of Human Factors in Housing*, vol. 3 (Stanford: Stanford University Press, 1950).

Freeman, Scott, Jeffrey Grogger, and Jon Sonstelie, "The Spatial Concentration of Crime," *Journal of Urban Economics* 40 (1996), 216–231.

Friendly, Michael, "A.-M. Guerry's Moral Statistics of France: Challenges for Multivariable Spatial Analysis," *Statistical Science* 22 (2007), 368–399.

Glaeser, Edward L., and Bruce Sacerdote, "Why Is There More Crime in Cities?" *Journal of Political Economy* 107 (1999), 225–258.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics* 111 (1996), 507–548.

Gottfredson, Michael R., and Travis Hirschi, *A General Theory of Crime* (Stanford: Stanford University Press, 1990).
Goudriaan, Heike, James P. Lynch, and Paul Nieuwbeerta, "Reporting to the Police in Western Nations: A Theoretical Analysis of the Effects of Social Context," *Justice Quarterly* 21 (2004), 933–969.
Hagestad, Gunhild O., and Peter Uhlenberg, "Should We Be Concerned about Age Segregation? Some Theoretical and Empirical Explorations," *Research on Aging* 28 (2006), 638–653.
Hausman, Jerry A., Jason Abrevaya, and Fiona M. Scott-Morton, "Misclassification of the Dependent Variable in a Discrete-Response Setting," *Journal of Econometrics* 87 (1998), 239–269.
Kleiber, Christian, and Achim Zeileis, *Applied Econometrics with R* (New York: Springer-Verlag, 2008).
Lee, Sokbae, "Endogeneity in Quantile Regression Models: A Control Function Approach," *Journal of Econometrics* 141 (2007), 1131–1158.
Lewbel, Arthur, "Identification of the Binary Choice Model with Misclassification," *Econometric Theory* 16 (2000), 603–609.
Manski, Charles F., "Identification of Binary Response Models," *Journal of the American Statistical Association* 83 (1988), 729–738.
——— "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60 (1993), 531–542.
——— "Economic Analysis of Social Interactions," *Journal of Economic Perspectives* 14 (2000), 115–136.
Mears, Daniel P., Matthew Ploeger, and Mark Warr, "Explaining the Gender Gap in Delinquency: Peer Influence and Moral Evaluations of Behavior," *Journal of Research in Crime and Delinquency* 35 (1998), 251–266.
Morselli, Carlo, and Pierre Tremblay, "Criminal Achievement, Offender Networks and the Benefits of Low Self-Control," *Criminology* 42 (2004), 773–804.
Nieuwbeerta, Paul, Patricia L. McCall, Henk Elffers, and Karin Wittebrood, "Neighborhood Characteristics and Individual Homicide Risks Effects of Social Cohesion, Confidence in the Police, and Socioeconomic Disadvantage," *Homicide Studies* 12 (2008), 90–116.
Patacchini, Eleonora, and Yves Zenou, "Juvenile Delinquency and Conformism," *Journal of Law, Economics, and Organization* 28 (2012), 1–31.
R Core Team, *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing, 2014).
Sampson, Robert J., Thomas Gannon-Rowley, and Jeffrey D. Morenoff, "Assessing 'Neighbourhood Effects': Social Processes and New Directions in Research," *Annual Review of Sociology* 28 (2002), 443–478.
Sampson, Robert J., and W. Byron Groves, "Community Structure and Crime: Testing Social Disorganization Theory," *American Journal of Sociology* 94 (1989), 773–802.
Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls, "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy," *Science* 277 (1997), 918–924.
Silverman, Dan, "Street Crime and Street Culture," *International Economic Review* 45 (2004), 761–786.
Steffensmeier, Darrell, and Emilie Allan, "Gender and Crime: Toward a Gendered Theory of Female Offending," *Annual Review of Sociology* 22 (1996), 459–487.
Sutherland, Edwin H., *Principles of Criminology* (Chicago: University of Chicago Press, 1947).
Sykes, Gresham M., and David Matza, "Techniques of Neutralization: A Theory of Delinquency," *American Sociological Review* 22 (1957), 664–670.
Walker, Joan L., Emily Ehlers, Ipsita Banerjee, and Elenna R. Dugundji, "Correcting for Endogeneity in Behavioral Choice Models with Social Influence Variables," *Transportation Research Part A: Policy and Practice* 45 (2011), 362–374.
Weerman, Frank M., "Delinquent Peers in Context: A Longitudinal Network Analysis of Selection and Influence Effects," *Criminology* 49 (2011), 253–286.
Wikström, Per-Olof H., Dietrich Oberwittler, Kyle Treiber, and Beth Hardie, *Breaking Rules: The Social and Situational Dynamics of Young People's Urban Crime* (Oxford: Oxford University Press, 2012).
Wilsem, Johan van, Karin Wittebrood, and Nan Dirk de Graaf, "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 (2006), 226–247.
Zenou, Yves, "The Spatial Aspects of Crime," *Journal of the European Economic Association* 1 (2003), 459–467.