

UC Irvine

UC Irvine Previously Published Works

Title

Extrapolative Simulation of Neighborhood Networks based on Population Spatial Distribution: Do they Predict Crime?

Permalink

<https://escholarship.org/uc/item/3b70651v>

Journal

Social Networks, 35(4)

Authors

Hipp, John R
Butts, Carter T
Acton, Ryan M
[et al.](#)

Publication Date

2013-12-10

Peer reviewed

Extrapolative Simulation of Neighborhood Networks based on Population Spatial

Distribution: Do they Predict Crime?

John R. Hipp*
Carter T. Butts
Ryan Acton
Nicholas N. Nagle
Adam Boessen

March 18, 2013

Post-print: published in Social Networks (2013) 35(4): 614-625

Word count: 9,035

Word count (including references): 11,059

Running Head: "Crime and simulated networks"

* Department of Criminology, Law and Society and Department of Sociology, University of California, Irvine. Address correspondence to John R. Hipp, Department of Criminology, Law and Society, University of California, Irvine, 3311 Social Ecology II, Irvine, CA 92697; email: john.hipp@UCI.edu. This research is supported in part by NSF grant BCS-0827027

Extrapolative Simulation of Neighborhood Networks based on Population Spatial

Distribution: Do they Predict Crime?

Abstract

Objectives: Previous criminological scholarship has posited that network ties among neighborhood residents may impact crime rates, but has done little to consider the specific ways in which network structure may enhance or inhibit criminal activity. A lack of data on social ties has arguably led to this state of affairs. We propose to avoid this limitation by demonstrating a novel approach of extrapolatively simulating network ties and constructing structural network measures to assess their effect on neighborhood crime rates.

Methods: We first spatially locate the households of a city into their constituent blocks. Then, we employ spatial interaction functions based on prior empirical work and simulate a network of social ties among these residents. From this simulated network, we compute network statistics that more appropriately capture the notions of cohesion and information diffusion that underlie theories of networks and crime.

Results: We show that these network statistics are robust predictors of the levels of crime in five separate cities (above standard controls) at the very micro geographic level of blocks and block groups.

Conclusions: We conclude by considering extensions of the approach that account for homophily in the formation of network ties.

Keywords: neighborhoods, crime, social networks, spatial effects, simulation

Bio

John R. Hipp is an Associate Professor in the department of Criminology, Law and Society, and Sociology, at the University of California Irvine. His research interests focus on how neighborhoods change over time, how that change both affects and is affected by neighborhood crime, and the role networks and institutions play in that change. He approaches these questions using quantitative methods as well as social network analysis. He has published substantive work in such journals as *American Sociological Review*, *Criminology*, *Social Forces*, *Social Problems*, *Mobilization*, *City & Community*, *Urban Studies* and *Journal of Urban Affairs*. He has published methodological work in such journals as *Sociological Methodology*, *Psychological Methods*, and *Structural Equation Modeling*.

Carter T. Butts is an Associate Professor in the department of Sociology and the Institute for Mathematical Behavioral Sciences at the University of California, Irvine. His research involves the application of mathematical and computational techniques to theoretical and methodological problems within the areas of social network analysis, mathematical sociology, quantitative methodology, and human judgment and decision making. His work has appeared in a range of journals, including *Science*, *Sociological Methodology*, the *Journal of Mathematical Sociology*, *Social Networks*, and *Computational and Mathematical Organization Theory*.

Nicholas N. Nagle is an Assistant Professor in the department of Geography at the University of Tennessee.

Ryan Acton is an Assistant Professor in the department of Sociology at the University of Massachusetts, Amherst.

Adam Boessen is a doctoral candidate in the department of Criminology, Law and Society at the University of California, Irvine. His primary research interests include the community of context of crime, spatial analysis, social networks, and juvenile delinquency. He uses quantitative methodologies to examine networks and neighborhood processes, the relationship between daily activities and crime, and the impact of incarceration on juvenile offenders. His work has been published in *Crime and Delinquency*, *Social Networks*, and *The Annals of the American Academy of Political and Social Science*

Extrapolative Simulation of Neighborhood Networks based on Population Spatial Distribution: Do they Predict Crime?

A recurrent theme in the ecological study of crime is the theorized importance of social networks among residents for crime reduction. These theories are generally built on the premise that the social interactions among residents allow them to band together to provide informal social control that can help to reduce the level of crime in a neighborhood (Bursik 1988; Sampson and Groves 1989; Shaw and McKay 1942). It is sometimes posited that the structure of network ties leads to neighborhood cohesion and a willingness to provide informal social control in response to deviant behavior.

Although some theories have pointed to the possible importance of social networks for impacting the ecology of crime, prior research has given little consideration in to the actual *structure* of such networks. This is arguably in a large part due to data limitations. Nonetheless, the consequence is that little consideration is given to the specific structural characteristics that enable phenomena such as cohesion or collective efficacy (the sense that one's fellow residents are willing and able to act towards the common interest) to arise (Sampson, Raudenbush, and Earls 1997). Instead, studies typically measure some approximation of the *number* of ties in the neighborhood. However, this ignores the rich tradition in the social network literature of assessing how various network structural characteristics impact such phenomena as information flow and cohesion—both important in various criminological theories of neighborhood crime. There is thus substantial potential to advance understanding of neighborhood crime processes by better exploiting the large body of knowledge in this area. Indeed, a few studies in the neighborhoods and crime literature have attempted to create proxies of the network structure by

measuring the reported strength of ties under the hypothesis that the presence of more weak ties is indicative of a more expansive network (Bellair 1997). Our study extends this literature by constructing explicit measures of the network structure.

The lack of available empirical data regarding these questions is due to the formidable data and cost challenges: assessing many network properties of theoretical interest requires a census of all persons in the city, as well as the spatial distribution of all of their ties, and therefore typical sampling-based survey strategies are not appropriate. Clearly, this is impractical for researchers.

The present study addresses this challenge with a novel approach: we *extrapolatively simulate* networks from census and other data sources, and employ properties of the simulated networks to make empirical predictions regarding crime rates. Recognizing that the existing criminological literature implicitly focuses almost exclusively on one social network measure (mean degree) in studying the effect of networks on crime rates, we will consider several social network measures that capture mechanisms predicted by criminological theory. Although our approach may appear somewhat unconventional at first glance, we argue below that numerous prior studies in the neighborhood and crime literature have implicitly used a very simple “simulation” model of social tie choice, and then considered only the network measure of density in testing the effect on crime rates.

Although accounting for all of the processes that generate social ties within a community is a potential future direction for the method we employ, we view the present study as a proof of concept. That is, our approach is that of an intellectual model, rather than an emulative model (Carley 2002; Carley 2009), as we build a very simple model and show its behavior in these settings rather than attempting to build a model that incorporates all information known about

neighborhood networks. We therefore specify a relatively simple tie formation model to demonstrate the utility of the approach. Given that prior criminological theory has only peripherally considered the role of propinquity in determining the structure of social ties (and therefore has underappreciated its importance), in this study we will focus explicitly on the influence of spatial distance on tie probability. We therefore build on prior research by Butts and colleagues (Butts, Acton, Hipp, and Nagle 2011) that assessed the extent to which simulating spatial networks based on a simple distance decay function could provide predictions about the macro structure of network ties in a community; here, we assess whether this network structure can also be used to explain the more micro phenomenon of crime events. Although this is a simplification of the true tie process, propinquity is nonetheless a well-known powerful mechanism for tie formation, and our simulations extrapolated from interaction and population data thus plausibly reflect the general contours of real urban networks. As we show, even this very broad approximation is sufficiently powerful to provide robust results for crime event prediction in five US cities, and therefore is an effective proof of concept of this general approach. Future work can extend this model, and the strategy in general, in various directions.

Crime in neighborhoods

The role of social ties in reducing crime

Although there are several key ecological theories of crime, each positing various mechanisms through which neighborhood demographics affect criminal behavior, a common theme running through many of these theories is the putative role of the network of social ties among residents. For example, routine activity theory (Cohen and Felson 1979) posits that the presence of guardians can reduce the incidence of crime, suggesting that residents in neighborhoods can play an important role in combating crime. Similarly, social disorganization

theory posits that social ties are important for enabling residents to employ informal social control to reduce the amount of crime in their neighborhoods (Bursik 1988; Sampson and Groves 1989; Shaw and McKay 1942). In some tests of this theory, it is posited that the mere presence of more positive social ties on average for the residents of a neighborhood (which is essentially *mean degree*) will in and of itself increase cohesion and therefore informal social control (without regard to the distribution of those ties, nor their configuration) (Bursik 1988; Sampson and Groves 1989; Warner and Rountree 1997). The importance of resident social networks was also discussed by Sampson (2006), who argued that social ties among residents may be a necessary precondition for the formation of *collective efficacy* in neighborhoods, which also impacts crime rates.

In these existing theoretical models of how neighborhood networks might impact crime, it is often posited that, A) certain characteristics of the residents in a neighborhood affect the formation of social ties, and B) these social ties enable the provision of informal social control, which, C) would then reduce the level of crime. Rarely do studies attempt to assess this entire model: instead, some studies focus on path A and try to predict the prevalence of ties among residents of the neighborhood (Bellair 1997; Campbell and Lee 1992; Sampson 1988; Sampson 1991; Woldoff 2002). These include studies outside of the criminology literature. Numerous studies in the social disorganization literature assess the effect of certain neighborhood structural characteristics—racial/ethnic heterogeneity, residential instability, and concentrated disadvantage--on crime rates (Hipp 2007b; Peterson, Krivo, and Harris 2000; Warner and Pierce 1993). Such studies are assessing the total effect of such structural characteristics on crime assuming that they indeed affect network ties (path A) and that network ties then affect informal social control (path B) which then reduces crime (path C). Of course, if paths B and C are

indeed accurate, but other structural characteristics also affect the formation of network ties (path A) other than those specified by social disorganization theory, then these other structural characteristics would also impact crime rates through such an indirect route. Our study, on the other hand, focuses on the combined paths B and C.

It is important to emphasize that although studies have focused on how crime is impacted by the characteristics of ties such as their strength (Sampson 2004), their frequency (Bellair 1997), or the race (Warner and Rountree 1997) or gender of those with more ties (Rountree and Warner 1999), studies rarely consider the *structure* of the overall network. That is, which specific structural network characteristics might be important for fostering informal social control behavior? The few studies in the criminological literature that attempt to “directly” measure the networks of residents typically ask respondents how many neighbors they know, and then compute the mean among residents in the neighborhood (which is, in fact, mean degree) (Rountree and Warner 1999; Warner and Rountree 1997). These studies are implicitly testing the total effect of paths B and C. Studies that test path B—whether social networks affect informal social control behavior—generally focus on a measure of mean degree (Bellair and Browning 2010; Warner 2007). We argue that there is little reason to suspect that mean degree is the only, or even the most appropriate, possible network measure that would be of importance for neighborhood crime rates. For example, there is a distinction between a) *private social control* (i.e., the activities of families on their members to sanction behavior); b) *parochial social control* (typically control activity conducted within a neighborhood); and c) *public social control* (i.e., obtaining resources from the larger community to address problems in the local neighborhood) (Bursik and Grasmick 1993). These different conceptualizations point to the importance of the larger network structure for impacting crime control, and some theorizing has

pointed to the possible importance of broader social network ties (Bellair 1997). We suggest that the lack of empirical data on complete networks among residents in cities has limited development of understanding the impact of the entire network structure.

In thinking about the role of social networks among residents, it is useful to begin with the question of *why* social ties might reduce the amount of crime in neighborhoods. We suggest that there are at least two mechanisms through which they might operate: 1) by increasing cohesion (both within the neighborhood, and across neighborhoods); and 2) by transmitting information. Each of these has implications for which structure of social ties is most beneficial to crime reduction. We consider each of these in turn.

Scholars in the social network literature have developed a number of structural measures intended to capture the degree of cohesion within a particular network, or a subgroup of a larger network. We focus here on three such measures that are each built on a different underlying theoretical model. One approach focuses on the number of ties among members of a defined subgroup, and therefore the average number of ingroup ties by members of the network or group (within group mean degree) is one measure of cohesion (Wasserman and Faust 1994).¹ Thus, groups with higher average number of ties among members are considered more cohesive.

A second approach uses the density of within-group ties as a measure of cohesion. Thus, the proportion of ties that exist out of all possible ties that might exist within a defined subgroup (Blau 1977; Lakon, Godette, and Hipp 2007). This measure changes the denominator from population (for mean degree) to, essentially, population squared. Thus, this captures not the number of ties per person, but rather the proportion of all possible ties, and as such expresses the notion that cohesion depends not only on group members to whom one is tied, but also those to

¹ Moody and White (2003) created a measure that explicitly defined cohesion based on the notion of the vulnerability of the group to the removal of one or a few of the actors.

whom one is *not tied*. Note that this differs from mean degree in that mean degree must increase with population size in order for density to be maintained: if mean degree is roughly fixed, increasing population will force density to decline. Under the spatial model, increasing population proportionately with the same geographical distribution maintains approximate density (instead increasing mean degree).

A third approach is to measure cohesion based on the ratio of within-group ties to out-group ties. Thus, network scholars frequently measure the cohesion of a subgroup based on the degree to which social ties are *within* group rather than *across* group (e.g. the LS sets of Luccio and Sami 1969). Whereas network researchers frequently consider endogenous notions of "group" in this regard, this same principle of internal versus external contact can also be applied to groups induced by exogenous characteristics (such as membership in a physical neighborhood). For instance, Krackhardt and Stern (1988) introduced a measure of cohesion known as the E-I index, which measures the extent to which ties associated with a pre-defined group are concentrated on in-group versus out-group members. By measuring the "inwardness" versus "outwardness" of a group's ties, Krackhardt and Stern were able to predict differences in the performance of organizational units subjected to adverse conditions. Under strain, highly inward-focused groups rallied internally, but were unable to cooperate with others to advance their collective interest; by contrast, more outwardly focused groups were better able to sustain cooperation with other units, and to thereby regain lost performance.

Moving beyond these three measures of cohesion for neighborhoods which might enhance *parochial control* (Hunter 1995; Taylor 1997), measures capturing the linkage of ties into the broader community may be required to capture the potential for *public control* (e.g., political action, influence over the policing process, coordination with other neighborhoods)

(Bellair 2000; Bursik and Grasmick 1993; Hunter 1995). A network measure that arguably captures such cohesion with the larger community is the average k -core (Seidman 1983) number. k -cores are defined as sets of vertices such that every member of the set is adjacent to at least k other members of the set.² An important property of k -cores is that they are locally cohesive: although cores need not be connected, every member of a high-order core belongs to a locally cohesive set. For $k > 1$, these sets necessarily contain spanning cycles, and are thus robust to the loss of edges or vertices.³ Those belonging to high-order cores have a dramatically enhanced chance of being able to reach some local set of individuals relative to those belonging to low-order cores. The average core number in a given population thus serves as a natural proxy of the extent to which members of the population are typically part of such robustly connected subgroups.

Beyond capturing cohesion, networks can also transmit *information* between residents. Information regarding criminal activity is a necessary precondition for residents to get involved in improving the neighborhood: clearly, residents must be aware of neighborhood problems in order to mobilize to address them. There is some evidence that networks might facilitate information flow about burglaries in neighborhoods (Warner and Rountree 1997). Not only does information about problems flow through interpersonal networks, but so too does information about fellow residents' willingness to engage in problem-solving behavior. The sense of collective efficacy derived from the knowledge that one's fellow residents are willing and able to act towards the common interest is arguably necessary to encourage residents to engage in behavior that otherwise may not benefit themselves directly (Sampson, Raudenbush, and Earls

² A k -core is a maximal group of actors in which all are tied to at least k other actors in the core. Larger values of k imply smaller groups, but with more connections.

³ To define these network terms: *vertices* are the persons in the sample; *adjacent* refers to the presence of a tie; *spanning cycles* are a cycle through a set of vertices (e.g., person A is tied to person B is tied to person C, etc).

1997).

There are numerous measures that capture the potential of a network to transmit information in the most efficient manner. The literature measuring *diffusion potential* has focused on the structure of the network (Moody 2002; Valente 1995; Valente and Davis 1999). A key insight from this perspective is that the capacity for information to flow throughout the network is potentially limited and unequally distributed across network members. In some network structures, information can flow easily to a large number of persons. On the other hand, some network structures inhibit the flow of information across groups, instead constraining it within particular subgroups. One approach considers how information randomly introduced into the network then can flow throughout the network given a particular structure (Butts 2010).

It is important to consider the geographic scale at which these networks likely affect cohesion and information flow (Hipp 2007a). The geographic scale at which cohesion operates is typically not very clearly specified in theories. Nonetheless, we argue that cohesion likely operates at a relatively small geographic scale. The fact that the density of ties among residents drops dramatically with increasing distance between residents suggests that a small unit such as a block might be most appropriate (Grannis 2009). We thus expect network cohesion measures to exhibit their strongest effects for small units such as blocks, and to be weaker for larger geographic aggregations.

Modeling social tie formation

A theoretical question is how social ties are formed in neighborhoods (path A from the earlier model described). Social network theory posits that *homophily* and *propinquity* are two key processes driving the formation of social ties. The notion of *homophily*—that is, the preference to form ties with others who are similar to oneself—is pervasive in the social network

literature (for a nice review, see McPherson, Smith-Lovin, and Cook 2001). Homophily can be induced by various social statuses, including race/ethnicity, economic resources, marital status, religious affiliation, the presence of children, etc. Thus, persons who are similar on these various characteristics are more likely to form and retain social ties. When considering the role of homophily, social disorganization theory focuses almost exclusively on the importance of homophily based on race/ethnicity (Sampson and Groves 1989; Veysey and Messner 1999). In practice, researchers often simply include a measure of racial/ethnic heterogeneity in the neighborhood and test whether it affects the crime rate, with the implicit, or sometimes explicit, mechanism being these social ties (Hipp 2007b; Roncek and Maier 1991; Sampson and Groves 1989; Warner and Rountree 1997). However, a consequence of ignoring the more general principle of social distance is that ecological studies of crime rarely consider other social dimensions beyond race/ethnicity. Exceptions are recent work considering the social distance engendered by economic inequality (Hipp and Perrin 2009), and a few studies testing whether the presence of economic inequality in neighborhoods is associated with higher crime rates (Crutchfield 1989; Hipp 2007b; Messner and Tardiff 1986). Another exception is a study measuring social distance simultaneously along a number of socio-demographic dimensions for the residents in a micro-neighborhood, and assessing the effect of this social distance on collective perceptions of disorder and crime (Hipp 2010).

A small but growing body of research in the social networks literature studies the effect of *propinquity* by focusing on the functional form of the spatial distribution of social ties of residents. This literature has shown a strong effect of spatial proximity on tie formation and that this effect can be captured by a distance decay function. This propinquity effect was demonstrated on the formation of ties in micro-areas in two early studies (Caplow and Forman

1950; Festinger, Schachter, and Back 1950), as well as more recent studies (Grannis 2009; Hipp and Perrin 2009). Although these studies focused on the network of ties among residents in relatively small geographic areas that are the size of neighborhoods or smaller and showed the greater likelihood of tie formation to those physically nearby, scholars have generalized these ideas and posited a distance decay function to capture the spatial distribution of ties over a much larger area (Butts Forthcoming).

Criminological scholars have not considered the consequences of a distance decay function for neighborhood network ties. Instead, the consideration of propinquity is generally limited to testing for a contextual effect in which higher overall neighborhood residential stability affects tie formation (Logan and Spitze 1994; Sampson 1988; Sampson 1991; Warner and Rountree 1997): this assumes that physical closeness leads to greater tie formation over time. Studies often test whether such residential stability is related to crime rates, under the assumption that it works through the mechanism of increased social ties (Bellair 2000; McNulty and Holloway 2000; Warner and Pierce 1993).

The current study

The present study builds upon prior research by Butts and colleagues (Butts, Acton, Hipp, and Nagle 2011) that adopted a similar simulation approach to derive macro predictions about the structure of community social networks. Butts et al developed a model of tie choice as a function of the spatial distance between residents based on the insight that social ties are inhibited by spatial distance between residents. This insight comes from the geography literature (see, e.g., Tobler 1970). Using a spatial interaction function that explicitly incorporates a distance decay, the social ties expected based on this model are simulated by using the actual geography of the households in the cities. We highlight that recent work by Butts (2010) has

shown that models based on distance effects alone can be quite robust to omitted tie formation mechanisms such as triad closure bias, even for complex phenomena such as diffusion potential. Using the same spatial interaction functions to simulate the networks, we extend this approach to five new cities. Importantly, we create measures of the structural properties of these networks and use them to predict the more micro consequence of actual crime rates in neighborhoods.

Our approach is unique in that we explicitly simulate the network of ties among residents. From the simulated network, we do not simply compute the mean degree of ties in a neighborhood, but instead construct several additional theoretically informed social network measures from these simulated networks that address the issues of cohesion and information flow that we identified above. We then employ these network measures to the prediction of crime rates in five U.S. cities. As we show, the network properties from the simulated networks are indeed able to predict the incidence of crime, above and beyond standard controls. Our findings thus provide both a novel vindication of the utility of our measures, and an elaboration that more completely addresses the connection between network structure and crime.

The present study adopts a relatively simple model of tie formation, and only focuses on impact of propinquity on tie formation; it can therefore be viewed as a “proof of concept” study. It is useful to assess if these network structures can predict a “real-world” phenomenon such as crime location. If using such a simple model of tie formation based on only physical distance can provide useful results, this implies that adding greater sophistication to the tie formation model will only improve the predictability of the model. Although it would seem natural to include homophily effects in the tie choice model, it should be emphasized that there is very little empirical evidence regarding the actual values of any homophily parameters that might be used for a simulation. Properly estimating such parameters requires information on both actual tie

choice, as well as the total choice set (e.g., who lives nearby). Existing literature almost never provides such estimates. Thus, introducing such covariates into the simulation would require adopting an unsatisfactory number of assumptions to come up with values for these covariates. In the present study, we instead followed the common strategy of simply including measures of racial/ethnic heterogeneity and inequality as crude proxies for homophily. Although it is also the case that various physical characteristics might impact the formation of social ties between residents (i.e., physical boundaries such as highways and rivers; the actual pattern of streets in a neighborhood), there is little evidence of the actual magnitude of such effects. Although we refer to the “network” among residents, it is well-known that numerous networks could be defined within a neighborhood based on various definitions of ties. Ties could be defined based on relatively strong or weak criteria, and the resulting network would differ based on these choices. However, better empirical evidence on what brings about strong versus weak ties in neighborhoods would be necessary to build a plausible simulation model. Our study is meant to demonstrate the utility of our approach, and we leave such complications for future research.

Data and Methods

Data

This study focuses on five cities for which we have data on the actual occurrence of reported crime events (“point data”). This allows us to place crime events from 2000-02 into the very small geographic container of blocks to capture possible micro-geography effects (Hipp 2007a; Weisburd, Bernasco, and Bruinsma 2009). We also aggregate to block groups to test more meso processes. The cities are Buffalo, Cincinnati, Cleveland, Sacramento, and Tucson. These cities provide variability across region of the country (3 eastern and 2 southwestern) and relative population density. We combine census data information on the location of households

in the blocks of these cities, along with GIS techniques to locate these blocks across the geography of the cities. The Census data provides information on the number of households in a block, along with the number of persons in each household. A limitation is that the upper bound category contains households of 9 or more persons, inducing a degree of uncertainty regarding the number of persons in these units; we place the extra persons randomly throughout the block.

With this information on the number of households in a block we generate points for these households and their members based on two different assumptions (since we do not have information on the actual location of the housing units). Under the first model, we assume a uniform micro-distribution; this implies a maximum entropy solution in which households are placed uniformly at random in the block. Under the second model, we assume a quasi-random micro-distribution; this implies a near-minimal entropy solution, in which households are placed in an extremely even manner using a low-discrepancy sequence (specifically, a two-dimensional Halton sequence). There is evidence that both approaches provide a reasonable approximation; given that the uniform approach appears to perform slightly better, we use it in all models (Almquist and Butts 2012).

To simulate the social networks, we need to employ a particular spatial interaction function (SIF) to model the geographic distribution of ties. The research in this area is sparse, making it difficult to know the functional form of the spatial effect of distance on tie formation. We estimated two different SIF's. The first we refer to as the "Festinger" SIF, as it is based on the Festinger, Schachter, and Back (1950) study in 1950 viewing the spatial distribution of a social friendship relation. This function is locally somewhat sparse, with a fairly long tail. This is of the general power law form with the SIF declining approximately with distance (d) raised to -2.8 . We also utilized a second SIF based on the Freeman et al. (1988) study of a face-to-face

interaction relation: this function is a locally dense relation that attenuates very quickly with distance, as the SIF declines approximately with d raised to -6.4 . Given that the general pattern of results were similar using both of these point placement patterns, and both of the SIF's, for brevity we present only those from the Festinger SIF using the quasi-random point distribution. A more complete description of the point placement and the simulation is provided in the Technical Appendix.

Dependent variables

Our outcome variables are counts of violent crime and property crime over the 2000-02 period. We combine aggravated assaults, robberies, and homicides into the count of *violent crime*. We combine burglaries, motor vehicle thefts, and larcenies into the *property crime* count. We estimated ancillary models with these crime types disaggregated and the pattern of results was similar to that presented here for these two aggregated measures. We therefore only present the violent and property crime models.

Independent variables

We included several network population distribution measures intended to capture the theoretical ideas described above. We included three social network measures that adopt different approaches to capture cohesion within the neighborhood. First, following prior research, we include *mean degree*. This is simply the average number of social ties for persons in the neighborhood. Second, we include a measure of *tie density*, given that this is frequently used as a measure of cohesion (Wasserman and Faust 1994). This measure divides the number of social ties in the neighborhood by the number of possible ties (which is $p(p-1)/2$, where p = population), and represents a theoretical model in which neighbors to whom one is untied negatively affect cohesion. The third measure captures the extent to which ties are local: the

proportion of ties that are within the geographic unit (measured at three geographic units: within the *block*, within the *block group*, and within the *tract*). We use these three units given that there is uncertainty of the proper unit that captures a “neighborhood”: blocks capture very local processes, block groups capture the meso processes of neighborhoods as we measure them here, whereas tracts are larger and prior research has often used them to measure “neighborhoods”. Larger values indicate a higher proportion of ties within the geographic unit.

We also created a measure to capture cohesion not bounded within the neighborhood. This measure of *k-cores* is the average of k^i within the neighborhood, where k^i is the order of the highest *k*-core to which the *i*th resident in the neighborhood belongs. A resident with a high core number belongs to relatively well-connected sets, while those with low core numbers belong to subgroups that are easily divided by the removal of nodes or edges; thus larger values for this measure indicate that residents are well-connected in the entire network of the city. Note that these *k*-cores are not constrained to the local neighborhood, but might also reach out into the broader community.

We created a measure of *diffusion potential* that likely captures a network structure that enhances the transmission of information. This measure was suggested by Butts (2010), and it captures the average size of the component to which each resident in the neighborhood belongs. This is measured as the maximum number of persons who could be reached by a signal introduced to an individual in the network; the expectation of this maximum is a natural measure of the diffusion potential of a network, with higher values indicating networks in which, on average, more members can be reached by randomly arriving information or resources.

To minimize the possibility of obtaining spurious results, we included several socio-demographic measures that are commonly included as predictors of crime rates in

neighborhoods. This first set of variables was included in models at all three geographic units of analysis. We take into account the racial/ethnic composition with measures of the *percent African American* and *Latino*. We also account for *racial/ethnic heterogeneity* with a measure of the Herfindahl index of five racial/ethnic groupings (white, African American, Latino, Asian, and other race), which takes the following form:

$$(1) \quad H = 1 - \sum_{j=1}^J G_j^2$$

where G represents the proportion of the population of ethnic group j out of J ethnic groups. To account for the presence of vacant units that might increase crime, we included a variable of the *percent occupied units* in the neighborhood. Given that homeowners may be more invested in the neighborhood and therefore engage in more crime-fighting behavior, we included a measure of the *percentage of units that are owned*. We included a measure of *population density*, in part because some criminologists suggest that high density areas will have more crime, and in part to test whether population density operates as a rough proxy for our network measures (given that the network is simulated based on the propinquity of ties). For the block-level models we constructed a measure of the *percent single parent households* (given that of the measures often included in an index of concentrated disadvantage, this is the only one available for blocks).

In the models for block groups, we were able to account for measures that are not available at the block level. Specifically, instead of the single parent household measure, we created a measure of *concentrated disadvantage* by combining four variables through principal components analysis: median household income, percent at or below 125% of the poverty level, percent single parent households, and the unemployment rate. We also included a measure of

income inequality in the neighborhood (based on the Gini coefficient).⁴

Table 1 displays the summary statistics for the network measures in the five cities broken out for blocks and block groups.⁵ As can be seen, the cities range in population density from those of the older, denser cities in the east of Buffalo, Cincinnati, and Cleveland, to the newer, more sprawling cities of the west in Sacramento and Tucson. Tucson is a particularly sprawling city. We see that mean degree is highest in Buffalo and Cleveland, and lowest in Tucson. However, diffusion potential follows a different pattern, as it is highest in Cleveland and Tucson, but lowest in Cincinnati. When viewing the proportion of ties within a geographic unit, the values are highest in the two low density cities (Sacramento and Tucson) and lowest in Cleveland. Thus, we can see that, in line with insights from urban sociologists, the idiosyncratic features of each urban environment are expected to produce regional network structures with rather different properties.

<<<Table 1 about here>>>

It is worth noting that even though our social networks are constructed entirely based on the principle of propinquity, the correlation between population density and our various network measures in general is not excessive. Only the measures of mean degree and k -cores are somewhat highly correlated with population density, and then only when aggregated to block

⁴ To account for the binned nature of the data, we use the prln04.exe program provided by Francois Nielsen at the following website: <http://www.unc.edu/~nielsen/data/data.htm>. This program accounts for binning by utilizing the Pareto-linear procedure Aigner, Dennis J. and Arthur S. Goldberger. 1970. "Estimation of Pareto's Law from Grouped Observations." *Journal of the American Statistical Association* 65:712-723, Kakwani, N. C. and N. Podder. 1976. "Efficient Estimation of the Lorenz Curve and Associated Inequality Measures From Grouped Observations." *Econometrica* 44:137-148., which Nielsen and Alderson adapted from the U.S. Census Bureau strategy Nielsen, Francois and Arthur S. Alderson. 1997. "The Kuznets Curve and the Great U-Turn: Income Inequality in U.S. Counties, 1970 to 1990." *American Sociological Review* 62:12-33..

⁵ Although we only display the results from the Festinger SIF using the quasi-random point placement, we briefly note that in the summary statistics, the "Festinger" SIF yields much larger personal networks than does the "Freeman" SIF, which is by design. The strong decay of the Freeman SIF results in much fewer social ties. On the other hand, the difference between using the quasi versus the uniform placement of household points has a much smaller impact on the size of these networks.

groups. For example, mean degree is correlated with population density .79 in block groups (but just .34 in blocks). The measure of tie density and the various measures of proportion within-geographic area ties have very little correlation with population density. The summary statistics for the outcome variables and control variables are presented in Table 2. The concentrated disadvantage factor score was created using the entire U.S., and thus these cities are nearly 1 standard deviation above the mean in disadvantage.

<<<Table 2 about here>>>

Methods

Given that our outcome variables are counts, we estimated negative binomial regression models (allowing us to account for overdispersion in these counts). To account for the varying size of the geographic units, we include the logged population size as an offset variable with a coefficient constrained to 1; this effectively transforms our outcome into a crime rate measure. We estimated fixed effects for the cities by including indicator variables for each city (with only five cities, a multilevel approach is not feasible).⁶

It is important to take into account the possibility that effects may differ based on the level of aggregation for ecological studies of crime (Hipp 2007a). We therefore estimated models from our five cities for two geographic units of analysis: blocks and block groups. For all models, we included all of the control variables described above. We first adopted a strategy of including our various network variables one at a time in models. After the initial models with these measures included separately, we estimated a model that included several network measures simultaneously, based on those shown to be important in the initial models.

Results

⁶ We also estimated models on each of the five cities separately to assess the robustness of the findings. In general, the pattern of results was quite robust over the cities, and we comment on this robustness throughout the results section.

Aggregating network measures to blocks

We turn first to the results for our network measures of cohesion. Mean degree does a good job explaining the location of crime when aggregating to blocks (the coefficient presented in the cell in the first row and column of Table 3 represents the coefficient from the *violent* crime models with block-level aggregations, when including all control variables). A one standard deviation increase in mean degree within a block reduces the violent crime rate 17.4% on average ($\exp(-.065*2.94=.826)$) and the property crime rate 23.6% ($\exp(-.092*2.94=.764)$).⁷

<<<Table 3 about here>>>

There is, however, little evidence that tie *density* reduces crime rates, despite its importance in other social network contexts. Tie density actually is associated with *higher* levels of violent and property crime (row 2 in Table 3), as a one standard deviation in tie density increases violent crime 18.7% and property crime 8.6%. Thus, whereas computing the number of ties as a proportion of persons (mean degree) shows evidence of lowering property and violent crime rates, computing them as a proportion of possible ties (tie density) results in higher property and violent crime rates. We thus infer that it is the geographic *concentration* of population that is more conducive to crime, holding constant the number of ties per person. Ties can be a resource for cohesion, but they must compete against the increasing opportunities for predation that arise within large, closely packed populations.

Arguably, the strongest effects are seen for the measures capturing cohesion as the proportion of ties within the local unit. In row 3 in Table 3, we see that in blocks in which resident ties are disproportionately to others in the same block there are much lower levels of violent crime, even when controlling for the typical measures included in ecological analyses of

⁷ The effect was quite robust across cities, as it was significant in all five cities in the property crime models, and in four of the five cities in the violent crime models (the only exception was Cincinnati).

crime. A one standard deviation increase in the proportion of ties within-block decreases the violent crime rate 20.6% on average. The effects are equally strong if we compute the tendency for local ties among block members but instead define the “in-group” to be either other residents in the same block group or tract (rows 4 and 5). We also see a strong protective effect on property crime rates when a higher proportion of ties are within-block. A one standard deviation increase in the proportion of ties within-block decreases the property crime rate 13.1% on average. When defining in-group ties based on the block group or the tract as the unit, a one standard deviation increase in the proportion of such ties reduce property crime rates nearly 20% (rows 4 and 5).

Turning to the measure capturing cohesion not bounded by the block, we see that higher values on the k -core measure (showing more robust local connectivity) are associated with lower crime rates (row 6). A one standard deviation increase in the measure is associated with 10% less violent crime on average and 19% less property crime.

Although the diffusion potential variable—meant to capture information flow—also has a negative effect on crime rates, it exhibits a weaker effect compared to the other network measures. A one standard deviation increase in the diffusion potential variable is associated with 8% less violent crime and 16% less property crime (row 7). This measure was a less consistent predictor in the models estimated on the cities one at a time.

Aggregating network measures to block groups

We next tested whether these network measures show similar effects when aggregated to the larger geographic unit of block groups rather than blocks. These results are presented in columns 3 and 4 of Table 3. The effect of the mean degree measure has now changed considerably. In row 1, mean degree measured at the block group is now associated with *higher*

levels of violent and property crime. A one standard deviation increase in mean degree in the block group is associated with 15% more violent crime and 8% more property crime, on average.⁸ Tie density continues to show a positive effect on these crime types at this larger aggregation, as a one standard deviation increase in tie density is associated with 17% more violent crime and 11% more property crime.

On the other hand, the measures of cohesion based on the proportion of within-group ties continue to exhibit robust negative relationships with crime rates. A one standard deviation increase in the proportion of ties within the same block group reduces the violent crime rate 17.6% on average. Furthermore, the clustering of ties within the blocks contained within the block group also reduces the violent crime rate an equal amount (row 3 in Table 4), whereas the clustering within the tract reduces it 10.8%. The story is similar, although somewhat weaker in the property crime models, as a one standard deviation increase in the proportion of within block ties among residents reduces property crime 9.6%, whereas a similar increase of within block group or within tract ties decreases it 15% and 7.6% respectively.

For the measure of cohesion not bounded at the neighborhood unit, the k -core measure now has a modest positive relationship with crime rates. The measure of information flow—the diffusion potential measure—now shows very little evidence that it captures areas with lower crime rates (row 7).

Including the network measures simultaneously: Aggregating network measures to blocks

We next turn to models including several of our network measures simultaneously. Along with the control variables, we included mean degree, the proportion of within area ties

⁸ It is interesting to note that this measure actually shows a modest negative effect on these crime rates in models estimated without the population density measure included. Although criminologists are mixed on the importance of including population density in such ecological models, we include it here to assess the extent to which our network measures based only on a spatial interaction can be approximated by population density.

based on three geographic units (blocks, block groups, and tracts), and the diffusion measure.

We were not able to include the measures of mean K-Cores and mean degree simultaneously in the model due to their very high degree of correlation (which ranged from .85 to .98 over various aggregation levels in various cities).⁹

The results of these models demonstrate that most of our findings when including the network measures separately hold up when including them simultaneously as seen in Table 4. The effect of block level mean degree is even more robust here when controlling for the proportion of within block ties, as well as these other measures. A one standard deviation increase in mean degree in the block is associated with 26% and 29% lower violent and property crime rates respectively, on average, holding constant the other measures in the model.

<<<Table 4 about here>>>

In these same models we also directly compared the effect of measuring the proportion of internal ties based on these three geographic definitions: blocks, block groups and tracts. We see that it is not just within-block ties that matter, but also within-block group and within-tract ties. A one standard deviation increase in the proportion of ties within the same block is associated on average with 22% less violent crime and 17% less property crime, whereas a one standard deviation increase in within-block group or within-tract ties reduce violent or property crime between 10 and 12%. Thus, we see that the presence of a higher proportion of more localized ties has a strong negative effect on the amount of violent crime, even when controlling for these other network measures, as well as the standard ecological measures that predict rates of crime. Furthermore, these in-group ties show a nesting effect in which there are additive effects for within block, block group, and tract ties in the same model: thus, a neighborhood that experiences a one standard deviation increase in the proportion of ties within the same block, the

⁹ The other variables in the model showed no evidence of collinearity, as all variance inflation factors were below 4.

same block group, and the same tract will experience 33% less property crime and 38% less violent crime.

The diffusion measure continues to show a negative effect on violent and property crime rates in the block models, although the effects are somewhat weaker than the models that did not simultaneously account for the other network measures. A one standard deviation increase in the diffusion potential measure is associated with 14.4% less violent crime and 15.9% less property crime.

Including the network measures simultaneously: Aggregating network measures to block groups

Turning to the models aggregating our measures to block groups, we again see that the only robust network measures are those capturing the importance of more localized ties. For both violent and property crime, the presence of more within-block group ties, and the presence of more intra-block ties within the block group, reduce crime rates. A one standard deviation increase in within-block ties reduces violent crime 13.8% and property crime 5.3%, whereas a one standard deviation increase in within-block group ties decreases violent crime 9.8% and property crime 13.6%.¹⁰ There is essentially no effect of mean degree in the block group in these models controlling for the other network measures. And whereas the diffusion potential measure does not impact violent crime, a one standard deviation increase in this measure reduces property crime 11.6%, when accounting for these other network measures.

It is important to emphasize how much the fit of these models is improved by including these network measures. Compared to models only containing the standard demographic

¹⁰ Viewing the models separately by city, it appears that whether within-block or within-block group ties is more important varies over city. For example, for the violent crime models, the presence of more intra-block ties within block groups is most important for reducing violent crime rates for all of the cities except for the city with the lowest population density (Tucson). For Tucson, it is instead the presence of more within block group ties that reduces violent crime rates. In the property crime models, although the presence of more within block group ties is most important for reducing property crime rates in Buffalo and Tucson, it is the presence of within block ties that are most important in the other three cities.

measures used in many criminological studies, adding these network measures to our full models improved the pseudo R^2 in the models aggregated to blocks 49% in the violent crime model and 91% in the property crime model (bottom row of Table 4). The model fit improvement was somewhat more modest in the models aggregated to block groups, as the pseudo R^2 was improved 13.5% for both crime types.

Conclusion

This study began by positing that important insights can be obtained by combining information on the actual geographic distribution of households in a community with the tendency of social ties to form based on spatial distance, and simulating the resultant social network of ties among residents. Even when specifying a very simple social interaction function in which residents form ties entirely based on propinquity, we showed that certain social network measures of cohesion show a strong negative relationship with the level of crime in the neighborhood. Of course, we do not know for certain whether such ties actually exist in these neighborhoods, and certainly the simulation would be wrong for some neighborhoods as it is a probabilistic model that captures general tendencies towards formation of social ties (although, as shown by Butts (2003; 2010), the baseline effects modeled here are extremely powerful predictors of structural properties under quite general conditions, and there is thus good reason to expect the models to be fairly robust). Despite those caveats, we observe that it would be difficult to explain away the predictive power of the network model without invoking some systematic bias that caused lower rates of crime to occur in the neighborhoods in which we also inaccurately simulated more cohesive networks. We know of no bias that would produce such results. Indeed, the robustness of the ability of putative network characteristics to predict crime above and beyond statistical controls -- and without requiring information on idiosyncratic

cultural or other characteristics of neighborhoods – are consistent with structural theories of crime inhibition. Although founded on the simple insight that the geographic distribution of persons can affect the formation of social ties—which may then affect crime rates—these strong and systematic results suggest exciting new theoretical possibilities for criminologists by incorporating more sophisticated models of tie formation. Including our social network measures considerably improved the model fit for both property and violent crime over both levels of aggregation.

Considering the three categories of effects examined here, we found that our measures of local cohesion from the simulated network showed a clear and interpretable pattern of results. Our results using these simulated ties showed that, although mean degree is consequential, the proportion of within-area ties was even more important for explaining areas with lower rates of crime. Furthermore, the importance of within-area ties generally held across the varying sized geographic units (blocks and block groups). This finding from the simulated networks implies that if studies were actually able to collect information on the full network of ties within a larger community, the presence of a higher proportion of these local ties would be important for detecting higher levels of perceived cohesion. Note that Sampson and Groves (1989) measured something similar in a sample based survey as they asked about the proportion of ties that are nearby. On the other hand, although tie density is a useful measure of cohesion in other contexts, it was not an important predictor of crime in these simulated networks.

There was also some evidence that our measure of cohesion *across* geographic units (*k*-cores) was associated with lower crime rates. Thus, blocks containing more persons linked into larger cores in the city generally had lower violent and property crime rates. This is consistent with the hypothesis that the provision of public control, as obtained from these far flung

networks, is important for reducing crime in the local neighborhood (Bellair 1997). This effect was only robust when aggregating to blocks, and was not present when aggregating to the larger units of block groups. Nonetheless, this measure was too highly correlated with mean degree for us to include it in the models containing multiple network measures.

The network measure we constructed to capture information flow showed a strong negative effect on crime rates in the models aggregated to small units (blocks), but a weaker effect in models aggregated to block groups. These were encouraging findings, but it is also possible that there are alternative network measures that would better facilitate the type of information flow that is important for crime fighting. Ours was simply an initial exploration into this approach, and leaves open the need for additional consideration of possible measures that might capture such external information. For example, one approach might compute the average number of blocks each resident in a block is linked to. This would assume that multiple ties to an external block would be redundant, and therefore the presence of a single tie to an external block would provide the necessary information. A twist on this would be to use the block as the unit of analysis, and compute the total number of external blocks the block is tied to (defined as any tie between a resident on the block and a resident in an external block). Here, the possible flow of information from an external block would enter the awareness of the block, and then could be transmitted among the block members. This might even imply an interaction between the presence of such external ties to blocks, and the density of within-area ties (which would be useful for then transmitting this information to fellow block members). Clearly, more theoretical consideration needs to be given to how social ties might convey useful information.

An important finding of the present study was that most of the observed effects were strongest when aggregating these simulated network measures to the very small geographic unit

of blocks. This implies that these are very micro-level processes (roughly on the order of a city block). When aggregating these measures to larger units, the relationships among variables frequently appeared considerably weaker or nonexistent. In particular, we note that the theoretical rationale given for these network effects -- processes such as the formation of cohesion -- is founded on phenomena that take place at small spatial scales. The coincidence of our effect strength pattern with the theorized effect scale further strengthens the argument that such micro-level mechanisms are indeed at work here. Although it is possible that other types of network features are predictive at these larger scales, such features would have to be motivated by theories other than those currently considered within the criminological literature. Of course, the fact that our control variables also appeared to perform much better when disaggregated to smaller units suggests that many of these ecological processes may operate at much smaller geographic units than is often posited in the criminology literature, suggesting the importance of carefully considering the appropriate unit of analysis (Hipp 2007a; Weisburd, Bernasco, and Bruinsma 2009).

It is worth pointing out that our results were also robust to two different point pattern placements and two different spatial interaction functions. That is, given that we only knew the number of housing units within blocks, and not their actual spatial location, we utilized two different approaches to placing these housing units in space. Likewise, although very little is known about the actual spatial interaction functions (SIF) of residents, we used two different SIFs from published research. The fact that our results were robust to these various specifications suggests that these may not be fragile findings, but may in fact be robust to various additional SIFs that we might specify. Nonetheless, future research will need to assess the degree to which this is indeed the case. For example, the spatial interaction function would

likely be different depending on the strength of the tie being measured.

Although we have highlighted some interesting possibilities in this line of research simulating networks based on structural properties, we emphasize that this is only an initial step in an exciting new line of research. An obvious next step is to incorporate homophily effects. Most prominent would be to take into account the fact that social tie formation is inhibited by racial/ethnic difference, economic difference, and age differences (Kalmijn and Vermunt 2007). Incorporating these covariate effects into the simulated networks would allow taking into account social distance, as well as physical distance (and various boundary effects from rivers and highways), for the formation of networks (Butts and Carley 2000; Hipp and Perrin 2009). A challenge to this additional direction is that there is little empirical evidence regarding the actual values of the parameters for such effects. Another limitation of this study was the focus on just five cities. Assessing whether simulated networks on additional cities behaved similarly for neighborhood crime rates would be useful. Nonetheless, it is worth emphasizing that our network measures generally did just as well, and frequently much better, than the typical measures included in models predicting the ecological distribution of crime. Furthermore, we also estimated the model separately on the five cities, and generally found quite robust results.

Another direction for future research would be to take into account that such networks not only inhibit crime, but that potential offenders are also part of the network. Criminologists have noted this possibility, speculating that the cohesion enabled by a network can also constrain activity by residents to the extent that they know family members of offenders (Browning, Feinberg, and Dietz 2004; Pattillo 1998). Even more provocative is the possibility that informationally "transparent" neighborhoods actually facilitate criminal activity, just as they facilitate its discovery. For instance, knowing when a neighbor will be out of town

simultaneously makes it easier to know when to watch for signs of trouble, and to identify the best time to break into his or her house. These considerations suggest the need to extend this simulation approach by incorporating information on potential offenders into the simulated networks. This also further suggests the need for more theoretical development in this area.

In conclusion, we believe our study has demonstrated an exciting new way for criminologists to think about how to integrate more formalized measures of social networks into studies of the ecological distribution of crime in cities. Rather than simply asking whether *more social ties* exist in a neighborhood, we have suggested that it is useful to consider the actual structure of the network of a community. The challenges of such data collection have arguably stunted any theoretical development in this direction. Our novel approach of specifying social networks from various propinquity preferences and the actual spatial distribution of households in the city, and then calculating various network structural measures based on the resultant social network suggests an exciting new direction for researchers. Future work incorporating other characteristics of tie formation and dissolution—including homophily preferences, as well as how mobility decisions affect the spatial distribution of ties—as well as other network structural measures that might capture important neighborhood characteristics, will allow further development of this approach.

References

- Aigner, Dennis J. and Arthur S. Goldberger. 1970. "Estimation of Pareto's Law from Grouped Observations." *Journal of the American Statistical Association* 65:712-723.
- Almquist, Zack W. and Carter T. Butts. 2012. "Point process models for household distributions within small areal units." *Demographic Research* 26:593-632.
- Bellair, Paul E. 1997. "Social Interaction and Community Crime: Examining the Importance of Neighbor Networks." *Criminology* 35:677-703.
- . 2000. "Informal Surveillance and Street Crime: A Complex Relationship." *Criminology* 38:137-167.
- Bellair, Paul E. and Christopher R. Browning. 2010. "Contemporary Disorganization Research: An Assessment and Further Test of the Systemic Model of Neighborhood Crime." *Journal of Research in Crime and Delinquency* 47:496-521.
- Blau, Peter M. 1977. "A Macrosociological Theory of Social Structure." *American Journal of Sociology* 83:26-54.
- Browning, Christopher R., Seth L. Feinberg, and Robert D. Dietz. 2004. "The Paradox of Social Organization: Networks, Collective Efficacy, and Violent Crime in Urban Neighborhoods." *Social Forces* 83:503-534.
- Bursik, Robert J. 1988. "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects." *Criminology* 26:519-551.
- Bursik, Robert J. and Harold G. Grasmick. 1993. *Neighborhoods and Crime*. New York: Lexington.
- Butts, Carter and Kathleen Carley. 2000. "Spatial Models of Large-Scale Interpersonal Networks." *Working paper*:1-47.
- Butts, Carter T. 2003. "Network inference, error, and informant (in)accuracy: a Bayesian approach." *Social Networks* 25:103-140.
- . 2010. "Bernoulli Graph Bounds for General Random Graphs."
- . Forthcoming. *Space and Structure: Models and Methods for Large-Scale Interpersonal Networks*, Edited by S. E. Fienberg and W. J. Linden. New York: Springer.
- Butts, Carter T., Ryan M. Acton, John R. Hipp, and Nicholas N. Nagle. 2011. "Geographical Variability and Network Structure." *Social Networks* 34:82-100.
- Campbell, Karen E. and Barrett A. Lee. 1992. "Sources of Personal Neighbor Networks: Social Integration, Need, or Time?" *Social Forces* 70:1077-1100.
- Caplow, Theodore and Robert Forman. 1950. "Neighborhood Interaction in a Homogeneous Community." *American Sociological Review* 15:357-366.
- Carley, Kathleen M. 2002. "Computational organizational science and organizational engineering." *Simulation Modelling Practice and Theory* 10:253-269.
- . 2009. "Computational modeling for reasoning about the social behavior of humans." *Computational and Mathematical Organization Theory* 15:47-59.
- Cohen, Lawrence E. and Marcus Felson. 1979. "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 44:588-608.
- Crutchfield, Robert D. 1989. "Labor Stratification and Violent Crime." *Social Forces* 68:489-512.
- Festinger, Leon, Stanley Schachter, and Kurt Back. 1950. *Social Pressures in Informal Groups*. Stanford, CA: Stanford University Press.
- Freeman, Linton C., Sue C. Freeman, and Alaina G. Michaelson. 1988. "On Human Social

- Intelligence." *Journal of Social and Biological Structures* 11:415-425.
- Grannis, Rick. 2009. *From the Ground Up: Translating Geography into Community through Neighbor Networks*. Princeton: Princeton.
- Hipp, John R. 2007a. "Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point." *American Sociological Review* 72:659-680.
- . 2007b. "Income Inequality, Race, and Place: Does the Distribution of Race and Class within Neighborhoods affect Crime Rates?" *Criminology* 45:665-697.
- . 2010. "Micro-structure in Micro-Neighborhoods: A New Social Distance Measure, and its Effect on Individual and Aggregated Perceptions of Crime and Disorder." *Social Networks* 32:148-159.
- Hipp, John R. and Andrew J. Perrin. 2009. "The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties." *City & Community* 8:5-25.
- Hunter, Albert. 1995. "Private, Parochial and Public Social Orders: The Problem of Crime and Incivility in Urban Communities." Pp. 204-225 in *Metropolis: Center and Symbol of Our Times*, edited by P. Kasinitz. New York: New York University.
- Kakwani, N. C. and N. Podder. 1976. "Efficient Estimation of the Lorenz Curve and Associated Inequality Measures From Grouped Observations." *Econometrica* 44:137-148.
- Kalmijn, Matthijs and Jeroen K. Vermunt. 2007. "Homogeneity of social networks by age and marital status: A multilevel analysis of ego-centered networks." *Social Networks* 29:25-43.
- Krackhardt, David and Robert Stern. 1988. "Informal Networks and Organizational Crises: An Experimental Simulation." *Social Psychology Quarterly* 51:123-140.
- Lakon, Cynthia M., Dionne C. Godette, and John R. Hipp. 2007. "Network-Based Approaches for Measuring Social Capital." Pp. 63-81 in *Social Capital and Health*, edited by I. Kawachi, S. V. Subramanian, and D. Kim. New York: Springer.
- Logan, John R. and Glenna D. Spitze. 1994. "Family Neighbors." *American Journal of Sociology* 100:453-476.
- Luccio, F. and M. Sami. 1969. "On the Decomposition of Networks in Minimally Interconnected Subnetworks." *IEEE Transactions on Circuit Theory* 16:184-188.
- McNulty, Thomas L. and Steven R. Holloway. 2000. "Race, Crime, and Public Housing in Atlanta: Testing a Conditional Effect Hypothesis." *Social Forces* 79:707-729.
- McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415-444.
- Messner, Steven F. and Kenneth Tardiff. 1986. "Economic Inequality and Levels of Homicide: An Analysis of Urban Neighborhoods." *Criminology* 24:297-317.
- Moody, James. 2002. "The Importance of Relationship Timing for Diffusion: Indirect Connectivity and STD Infection Risk." *Social Forces* 81:25-56.
- Nielsen, Francois and Arthur S. Alderson. 1997. "The Kuznets Curve and the Great U-Turn: Income Inequality in U.S. Counties, 1970 to 1990." *American Sociological Review* 62:12-33.
- Pattillo, Mary E. 1998. "Sweet mothers and gangbangers: Managing crime in a black middle-class neighborhood." *Social Forces* 76:747-774.
- Peterson, Ruth D., Lauren J. Krivo, and Mark A. Harris. 2000. "Disadvantage and Neighborhood Violent Crime: Do Local Institutions Matter?" *Journal of Research in Crime and Delinquency* 37:31-63.
- Roncek, Dennis W. and Pamela A. Maier. 1991. "Bars, Blocks, and Crimes Revisited: Linking

- the Theory of Routine Activities to the Empiricism of 'Hot Spots'." *Criminology* 29:725-753.
- Rountree, Pamela Wilcox and Barbara D Warner. 1999. "Social ties and crime: Is the relationship gendered?" *Criminology* 37:789-813.
- Sampson, Robert J. 1988. "Local Friendship Ties and Community Attachment in Mass Society: A Multilevel Systemic Model." *American Sociological Review* 53:766-779.
- . 1991. "Linking the Micro- and Macrolevel Dimensions of Community Social Organization." *Social Forces* 70:43-64.
- . 2004. "Networks and Neighbourhoods: The Implications of Connectivity for Thinking About Crime in the Modern City." Pp. 157-166 in *Network Logic: Who Governs in an Interconnected World?*, edited by H. McCarthy, P. Miller, and P. Skidmore. London: Demos.
- . 2006. "How Does Community Context Matter? Social Mechanisms, and the Explanation of Crime Rates." Pp. 31-60 in *The Explanation of Crime: Context, Mechanisms, and Development*, edited by P.-O. H. Wikstrom and R. J. Sampson. New York: Cambridge University Press.
- Sampson, Robert J. and W. Byron Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 94:774-802.
- Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918-924.
- Seidman, Stephen B. 1983. "LS sets as cohesive subsets of graphs and hypergraphs." *Mathematical Social Sciences* 6:87-91.
- Shaw, Clifford and Henry D. McKay. 1942. *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Taylor, Ralph B. 1997. "Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization." *Journal of Research in Crime and Delinquency* 34:113-155.
- Tobler, Waldo R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46:234-240.
- Valente, Thomas W. 1995. *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton.
- Valente, Thomas W. and Rebecca L. Davis. 1999. "Accelerating the Diffusion of Innovations Using Opinion Leaders." *ANNALS, AAPSS* 566:55-67.
- Veysey, Bonita M. and Steven F. Messner. 1999. "Further Testing of Social Disorganization Theory: An Elaboration of Sampson and Groves's 'Community Structure and Crime'." *Journal of Research in Crime and Delinquency* 36:156-174.
- Warner, Barbara D. 2007. "Directly Intervene or Call the Authorities? A Study of Forms of Neighborhood Social Control Within a Social Disorganization Framework." *Criminology* 45:99.
- Warner, Barbara D. and Glenn L. Pierce. 1993. "Reexamining Social Disorganization Theory Using Calls to the Police as a Measure of Crime." *Criminology* 31:493-517.
- Warner, Barbara D. and Pamela Wilcox Rountree. 1997. "Local Social Ties in a Community and Crime Model: Questioning the Systemic Nature of Informal Social Control." *Social Problems* 44:520-536.
- Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge.

- Weisburd, David, Wim Bernasco, and Gerben Bruinsma. 2009. "Putting Crime in Its Place: Units of Analysis in Spatial Crime Research." Pp. 256. New York: Springer Verlaag.
- Woldoff, Rachael A. 2002. "The Effects of Local Stressors on Neighborhood Attachment." *Social Forces* 81:87-116.

Tables

Table 1. Summary statistics for the network measures in five cities, using the quasi Festinger point pattern

	Buffalo		Cincinnati		Cleveland		Sacramento		Tucson	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Blocks										
Tie density	0.122	0.218	0.219	0.515	0.178	0.390	0.126	0.202	0.138	0.230
Mean degree	7.238	3.358	5.940	3.241	7.337	3.250	5.748	2.606	4.649	2.195
Proportion ties within block	0.490	0.147	0.483	0.194	0.444	0.153	0.491	0.153	0.499	0.166
Proportion ties within block group	0.768	0.139	0.791	0.155	0.760	0.144	0.852	0.132	0.829	0.136
Proportion ties within tract	0.911	0.112	0.889	0.130	0.859	0.136	0.931	0.101	0.924	0.105
K-cores	4.284	1.680	3.549	1.664	4.377	1.629	3.434	1.329	2.845	1.139
Diffusion potential	21.235	7.572	10.661	4.709	44.790	6.565	23.865	9.242	39.550	9.418
Population density (in 1,000's)	4.699	4.535	5.282	55.868	4.326	5.061	3.801	21.327	2.669	4.448
<i>N</i>	2,279		2,254		4,155		3,872		4,361	

Crime and simulated networks

Block groups											
Tie density	0.016	0.037	0.013	0.057	0.014	0.050	0.011	0.057	0.006	0.004	
Mean degree	7.658	2.842	6.236	2.881	8.005	2.817	6.102	2.539	5.315	2.116	
Proportion ties within block	0.539	0.102	0.596	0.131	0.508	0.110	0.579	0.130	0.594	0.116	
Proportion ties within block group	0.764	0.096	0.797	0.101	0.755	0.100	0.852	0.084	0.825	0.080	
Proportion ties within tract	0.912	0.071	0.892	0.085	0.864	0.094	0.939	0.053	0.924	0.059	
K-cores	4.226	1.498	3.500	1.400	4.432	1.454	3.325	1.179	2.896	0.896	
Diffusion potential	20.909	7.323	10.927	4.143	44.432	6.050	22.984	9.494	40.044	6.129	
Population density (in 1,000's)	4.256	2.931	3.172	2.517	3.908	1.837	2.826	1.385	2.239	1.086	
<i>N</i>	389		314		519		313		383		
<i>Note: population density measured in 1,000's of persons per square mile</i>											

Table 2. Summary statistics of variables used in the analyses at two units of analysis: blocks and block groups

	Blocks		Block groups	
	Mean	SD	Mean	SD
Outcome variables				
Violent crime events	2.74	5.91	26.12	27.48
Property crime events	15.80	37.17	151.87	142.97
Network variables				
Tie density	0.093	0.068	0.009	0.008
Mean degree	6.49	2.94	6.90	2.81
Proportion ties within block	0.50	0.13	0.49	0.10
Proportion ties within block group	0.81	0.13	0.78	0.10
Proportion ties within tract	0.91	0.10	0.90	0.08
K-cores	3.84	1.49	3.80	1.41
Diffusion potential	31.67	14.04	30.09	14.22
Independent variables				
Population density (in 1,000's)	0.43	2.55	0.41	0.41
Percent black	25.99	34.70	31.12	36.23
Percent Latino	16.88	22.32	13.96	19.23
Racial/ethnic heterogeneity	0.36	0.22	0.37	0.22
Percent single parent households	20.27	15.86		
Concentrated disadvantage			0.92	1.21
Inequality (Gini)			42.33	7.67
Percent occupied units	91.51	9.53	88.90	9.19
Percent owners	57.25	28.52	49.08	24.06
N				

Note: Cities are Buffalo, Cincinnati, Cleveland, Sacramento, and Tucson

Table 3. Models with violent or property crime as the outcomes, city fixed effects models. Units of analysis are blocks and block groups. Effects of various network measures

	Blocks				Block groups			
	Violent crime		Property crime		Violent crime		Property crime	
(1) Mean degree	-0.065 ***	-0.092 ***	0.050 ***	0.027 *	(-13.48)	(-29.89)	(3.54)	(2.53)
(2) Tie density	2.539 ***	1.217 ***	19.660 ***	13.558 ***	(15.27)	(10.34)	(5.70)	(4.61)
(3) Proportion of ties within block	-1.738 ***	-1.057 ***	-1.852 ***	-0.970 ***	(-21.86)	(-19.35)	(-11.98)	(-7.60)
(4) Proportion of ties within block group	-2.033 ***	-1.682 ***	-2.072 ***	-1.637 ***	(-22.66)	(-27.24)	(-10.28)	(-9.77)
(5) Proportion of ties within tract	-2.266 ***	-1.913 ***	-1.425 ***	-0.980 ***	(-21.49)	(-25.60)	(-5.56)	(-4.56)
(6) K-cores	-0.071 ***	-0.141 ***	0.105 ***	0.036 *	(-6.86)	(-22.53)	(4.31)	(1.97)
(7) Diffusion potential	-0.006 ***	-0.013 ***	0.007 *	-0.002	(-3.41)	(-10.08)	(2.27)	(-0.63)

*** $p < .001$ (two-tail test), ** $p < .01$ (two-tail test), * $p < .05$ (two-tail test). *T-values in parentheses. All models include all control variables described in the text for a particular unit of analysis, an intercept, and indicators for all but one cities. Logged population size is used as an offset variable in each negative binomial regression.*

Table 4. Models with violent or property crime as the outcomes. Units of analysis are blocks and block groups. Including network measures simultaneously in the model.

	Blocks				Block groups			
	Violent crime		Property crime		Violent crime		Property crime	
Diffusion potential	-0.011 ***		-0.012 ***		-0.004		-0.009 ***	
	(-6.40)		(-10.30)		(-1.42)		(-3.42)	
Mean degree	-0.102 ***		-0.117 ***		-0.002		-0.003	
	(-21.75)		(-36.56)		(-0.11)		(-0.26)	
Proportion ties within block	-1.857 ***		-1.371 ***		-1.423 ***		-0.521 **	
	(-19.52)		(-20.93)		(-6.90)		(-2.99)	
Proportion ties within block group	-0.884 ***		-0.977 ***		-1.046 ***		-1.483 ***	
	(-7.53)		(-12.20)		(-3.70)		(-6.19)	
Proportion ties within tract	-1.144 ***		-0.942 ***		-0.005		0.283	
	(-9.14)		(-10.66)		(-0.02)		(1.10)	
Percent black	0.008 ***		0.001 ***		0.005 ***		-0.001	
	(20.21)		(4.46)		(7.93)		(-1.36)	
Percent Latino	0.012 ***		0.005 ***		0.006 ***		-0.001	
	(20.31)		(11.16)		(5.32)		(-0.68)	
Racial/ethnic heterogeneity	0.631 ***		0.445 ***		0.004 ***		0.002 *	
	(10.67)		(11.00)		(4.64)		(2.02)	

Crime and simulated networks

Percent single parent households / concentrated disadvantage	0.008 *** (9.14)	0.003 *** (4.95)	0.256 *** (11.03)	0.057 ** (2.92)
Inequality (Gini)			0.001 (0.25)	0.003 (1.33)
Percent occupied units	-0.019 *** (-15.85)	-0.011 *** (-12.58)	-0.022 *** (-9.48)	-0.017 *** (-8.90)
Percent owners	-0.014 *** (-31.13)	-0.011 *** (-34.13)	-0.009 *** (-8.20)	-0.010 *** (-11.13)
Population density (in 1,000's)	-0.067 ** (-2.88)	-0.050 *** (-3.74)	-1.048 *** (-5.69)	-1.101 *** (-7.82)
Intercept	1.447 *** (8.61)	3.243 *** (27.01)	0.098 (0.27)	2.432 *** (7.92)
N	16,824	16,824	1,905	1,905
Pseudo R-square	0.076	0.052	0.079	0.043
Pseudo R-square (model with no network measures)	0.051	0.027	0.070	0.038
Percentage increase in variance explained for model with network variables compared to model without network variables	49.2%	91.5%	13.5%	13.5%
*** $p < .001$ (two-tail test), ** $p < .01$ (two-tail test), * $p < .05$ (two-tail test). T-values in parentheses. All models include indicators for all but one cities. Logged population size is used as an offset variable in each negative binomial regression.				

Technical Appendix

We refer to the spatial distribution of individuals within blocks as the population microdistribution. To simulate the microdistribution, we utilized two competing models representing the extreme cases of a range of potential point processes. The first (the uniform) assumes a maximum entropy solution, in which households (or isolated individuals) are placed uniformly at random subject to known geographical constraints. The second (the quasi-random) assumes a near-minimal entropy solution, in which households are placed in an extremely even, grid-like manner using a low-discrepancy sequence (specifically, a two-dimensional Halton sequence). By examining the impact of these extreme cases on social structure, we can thus infer the likely range of possibilities which could be occupied by processes having intermediate behavior. For both models, we account for ground-level congestion with a simple artificial elevation model to simulate multi-story residential structures in densely populated blocks. Specifically, households whose ground position would place them within a 10m radius of k previously placed households are given a vertical elevation of $4k$ meters; thus, intuitively, artificial elevation arises as population density grows, stacking new households on old ones. We treat arrival order as random. Finally, within-household proximity is maintained by requiring household size to satisfy the known marginals within each block, and then placing individuals at their household locations (jittering randomly within a 5m radius to avoid exact overlap).

While many choices of SIF are possible, we employ two specific functions obtained in prior work by Butts (2002) based on analyses of existing data sources. We take these as plausible examples of the types of SIFs one is likely to see from proximate relations such as those from which the functions were derived. The network is simulated as:

$$PR(Y = y | D) = \prod_{\{i,j\}} B(Y_{ij} = y_{ij} | \mathfrak{S}(D_{ij}, \theta))$$

where Y is the (random) graph adjacency matrix, D is a matrix of inter-vertex distances, B is the Bernoulli probability mass function (pmf), and \mathfrak{S} is the spatial interaction function (SIF) taking distances into the $[0; 1]$ interval (parameterized by real vector θ). The SIF provides the marginal probability of a tie between two randomly selected individuals at some given distance. Butts (2006a) has shown that the spatial Bernoulli graphs can be written as a special case of a more general curved exponential family of graph distributions. Given its simple interpretation and theoretical leverage, we employ the spatial Bernoulli framework here.

In computing distances for purposes of the SIF, we employ the Euclidean distances between individual surface positions in the projected geometry, plus a small correction to account for the effects of the built environment (where applicable). Specifically, differences in artificial elevation are added to the base Euclidean distance for those within 25m of one another, while the sums of individuals' artificial elevation distances are added for pairs whose surface positions differ by more than 25m. This simulates a basic feature of travel within the built environment, namely movement within a building for those who are otherwise proximate, versus movement to ground level, over to the second position, and up for those with distant surface locations. While one could employ more complex schemes (including explicit adjustment for roadways, local obstructions, etc.), this would require more detailed knowledge of household position, built environment, and other aspects of local geography than were available for our test regions. As a practical matter, experimentation with reasonable alternatives to the approach employed here did not produce substantively different results. By way of explanation, it should be noted that different notions of distance tend to be very strongly correlated, even at fairly small

scales. For instance, a comparison of Euclidean distances for the cases used here with Manhattan distances (the so-called city block metric sometimes suggested as an alternative for urban environments) produced a median correlation of approximately 0.99 under both uniform and quasi-random microdistribution models; even considering only very proximate points (within 100m of each other in Euclidean space), the median correlation is still approximately 0.98 for both microdistribution models. While the impact of alternative distance models on network structure at smaller spatial scales is an interesting and potentially productive target for further research, our experiments suggest that the results reported here are not sensitive to reasonable variations in how distance is defined.