# Close-ups and the scale of ecology: Land uses and the geography of social context and crime

Adam Boessen*
Department of Criminology and Criminal Justice
University of Missouri – St. Louis

John R. Hipp
Department of Criminology, Law and Society
University of California - Irvine

March 5th, 2015

Running Head: "The scale of ecology"

*Post-print. Published in Criminology 2015 53(3): 399-426*

* Address correspondence to Adam Boessen, Department of Criminology and Criminal Justice, University of Missouri – St. Louis, 531 Lucas Hall, One University Blvd., St. Louis, MO 63121. Email: boessena@umsl.edu This research is supported in part by NSF grant BCS-0827027.

# Close-ups and the scale of ecology:

# Land uses and the geography of social context and crime


## Abstract

Whereas one line of recent neighborhood research has placed an emphasis on zooming into smaller and smaller units of analysis such as street blocks, another line of research suggests that even the meso-area of neighborhoods is too narrow and that the area surrounding the neighborhood is also important. Thus, there is a need to examine the scale at which the social ecology impacts crime. We use data from seven cities from around the 2000-decade to test our research questions. Our results suggest that although many neighborhood factors appear to operate on the micro scale of blocks, others appear to have a much broader impact. In addition, we find that racially/ethnically homogenous blocks within heterogeneous block groups have the most crime. Our findings also show the strongest results for a multitude of land use measures and that these measures sharpen some of the associations from social characteristics. Thus, we find that accounting for multiple scales simultaneously is important in ecological studies of crime.


***Keywords***: neighborhoods, crime, aggregation, spatial effects

**Bio**

**Adam Boessen** is an Assistant Professor in the department of Criminology and Criminal Justice at the University of Missouri, St. Louis. His primary research interests include the community of context of crime, geography and space, and social networks.

**John R. Hipp** is a 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.

# Close-ups and the scale of ecology:

# Land uses and the geography of social context and crime

Building on the seminal work of the Chicago School in the early 20th century, a substantial body of literature has examined the ecology of crime. These studies often focus on the potential of social control in geographic areas and most frequently use “neighborhoods” as units of analysis. One well-known challenge in ecological studies is choosing a unit of analysis (Hipp 2007). In an effort to minimize within unit heterogeneity that may result from using larger units such as tracts (or groups of tracts), recent scholarship has suggested that the ecology of crime is best captured using a micro spatial scale/unit of analysis, including street blocks, street segments, or hot spots (Sherman, Gartin, and Buerger, 1989; Weisburd, Groff, and Yang 2012) . As this line of neighborhood research over the last decade has drilled down to smaller and smaller units, one risk is that researchers may adopt too narrow of a geographic lens and therefore miss important processes that occur at a broader spatial scale (or at least outside of one street block). For example, the ethnic heterogeneity of a street block may provide too narrow of a lens and might miss broader patterns in the surrounding area. Studies of micro-areas rarely simultaneously consider the more meso scale of neighborhoods or the social context surrounding these micro units.

Another body of research has focused on the meso-geographic scale of “neighborhoods” as “urban villages”. Two particularly strong assumptions of this approach are that social processes that produce crime are entirely contained within a neighborhood, and that the amount of crime is homogeneous across the smaller units within the neighborhood. Studies commonly include various measures of spatial processes to take into account how nearby neighborhoods might affect the level of crime in a focal neighborhood, and almost always find evidence of some type of spatial effect (e.g., see Mears and Bhati, 2006). This is all to suggest that a "neighborhood" as a unit of analysis appears unsatisfactory both because it is too *large*, and because it is too *small*.

Understanding the spatial scale of social dynamics is important not only when measuring socio-demographic characteristics posited to impact crime, but also when assessing the effects of physical characteristics. The environmental crime literature posits that land use features impact levels of crime by creating opportunities and situating where guardians might provide informal social control (Brantingham and Brantingham 1984). The extant research in this area usually only focuses on one land use characteristic in the local environment. Nonetheless, it is unclear whether such effects are very locally situated (at the actual location) or whether they have a broader spatial impact.

Thus, there is a need to assess the varying scales of ecological processes that produce crime, and in this paper, we begin by highlighting the different scales used in the ecology of crime literature to measure neighborhood processes. We begin to address this need here by incorporating crime and land use data at the block level for seven cities, and then computing ecological measures at three geographic units of analysis: 1) the local block; 2) the meso-neighborhood (the block group or tract); 3) the 5 miles surrounding a neighborhood with a distance decay. We assess the relative impact of these measures on various types of violent and property crimes. Our results suggest that different neighborhood processes do not all operate on the same scale as routinely assumed in the literature, and different conclusions are possible depending on the scale of analysis. The findings suggest that different crime processes may simultaneously operate at different spatial scales, and we also show that land uses are particularly important for understanding crime patterns.

## Social ecology of crime

*The process within neighborhoods*

A key theme of several criminological theories is that residents can provide social control through guardianship, and that this can impact levels of crime in the environment. This idea is present most prominently in social disorganization theory, but also exists in routine activities theory. Social disorganization theory posits that certain socio-demographic neighborhood compositions enhance the possibility of crime inhibiting behavior on the part of residents (Sampson and Groves 1989; Bursik 1988). Neighborhoods with more social interaction are expected to have more cohesion, and hence more willingness to confront offenders and others engaging in disorderly behavior. For example, the presence of racial/ethnic heterogeneity, residential instability, or concentrated disadvantage are posited to reduce the degree of interaction among residents, and hence the willingness to engage in social control behavior that might reduce the possibility of crime events.

Less clear in the social control literature is the proper scale at which to measure the potential for social control, the perceptions of residents, the social ties or sense of cohesion among residents. The critical questions are: How large of a spatial area should social control encompass? What is the spatial distribution of the social process of interest? Research asking residents about their perceptions regarding the size of their neighborhood typically find considerable differences across residents, including those living near one another. A study of Los Angeles found that whereas 36 percent felt their neighborhood was just their block, another 24 percent felt it was several blocks, 27 percent felt it was a fifteen minute walk, and 13 percent felt it was more than a 15 minute walk (Sastry, Narayan, Pebley, and Zonta, 2002). If the process of interest is simply where someone is willing to intervene to stop a criminal act, this implies a quite small spatial scale in that a person will need to be *visually* aware of the crime in order to stop it. On the other hand, perceptions of cohesion arguably have a much broader spatial scale in part because many neighborhood organizations are not strictly accessible to micro areas and many ties extend outside of the local area.

Nonetheless, most of the work in this area of research has defined the unit of interest as the "neighborhood", and therefore focused on more meso-level geographic units (e.g., Krivo and Peterson 1996; Hipp 2007; Bellair 1997). These studies often measure the demographic composition of these meso-units (under the assumption that certain demographic characteristics affect the level of informal social control behavior), and test whether this is associated with crime rates. Similarly, the vast majority of research on social disorganization and collective efficacy confines the focus of the social process within the neighborhood and does not include any measure of the nearby area (i.e., a spatial lag). Research typically posits, at least implicitly, that social control potential and/or behavior is best measured at the meso-level. However, the degree to which this is actually the case is rarely empirically examined.

Most often in neighborhood research a particular neighborhood unit is selected (e.g., Census tracts), and the process within the unit is expected to be wholly contained and uniform, regardless of the physical or social environment (Lee et al., 2008). Although there is no gold standard for a particular neighborhood unit, the social process of interest is always tied to some spatial area. Lee and colleagues' work on segregation indices suggests that the different social processes vary in spatial scale (see also Taylor 2015). One study examining the effect of neighborhood characteristics on prison misconduct included different neighborhood measures at different spatial scales (Boessen and Cauffman 2014). In their paper, a more micro measure (i.e., block groups) was included for residential instability, while a more macro measure (i.e, tracts) was used to capture racial/ethnic heterogeneity in the same model.

Whereas some scholars have suggested measuring ecological constructs at different scales (Hipp 2007; Taylor 2015), a necessary next step in this line of research is to *simultaneously* account for within unit heterogeneity, and not simply vary the size of units for different covariates. The ecology of communities suggests a need to consider a multiscale approach to better understand the crime process. The simultaneous consideration of multiple scales allows for a more complete and interdependent understanding of social processes by incorporating more information on other parts of the social system. Rather than relying exclusively on the social characteristics of neighborhoods, we also go beyond much prior research with the incorporation of land use data, which we will discuss later.

*The importance of micro-environments*

A growing number of scholars have suggested that it may be more appropriate to study crime at a much more micro scale (e.g., see Weisburd et al., 2012). In this perspective, the street block, street segment, or hot spot is the more appropriate ecological unit of analysis. In part, this might be because such small geographic units capture a more appropriate geographic scale at which much social interaction actually takes place (Grannis 2009). To the extent that ties foster cohesion and therefore a tendency to engage in informal social control behavior, the characteristics of micro-units would be more important to measure than would the larger meso-scale of neighborhoods. For example, a study that simulated social ties for a city found the strongest effects on crime for structural social network measures constructed at the block level, but weaker effects for larger aggregations such as block groups and tracts (Hipp et al., 2013).

Indeed, a burgeoning literature has focused on street blocks (or segments) as a unit of analysis when measuring the ecology of crime. These studies have typically shown that crime within cities disproportionately occurs on only a small proportion of blocks within the city (Weisburd et al., 2012), and this might be indicative of a hot spot (Sherman, Gartin, and Buerger, 1989). Such findings are often interpreted to indicate that street blocks are the ideal unit of analysis for studying the ecology of crime. One possibility, however, is that crime tends to cluster on a small number of street blocks due to the land use characteristics of those blocks, and not because of a lack of informal social control or guardianship, an issue to which we will return.

Neighborhood processes that are rooted in familiarity and proximity to neighbors are more likely at a smaller spatial scale. At small spatial scales, it is much more likely that residents will be more *familiar* with others, particularly those closer in geographic space, and thus more likely to form and maintain social ties, a stronger sense of cohesion, enhanced potential for information flow, and more potential for spatially induced resources (e.g., physical help from a nearby neighbor) (Hipp and Boessen, 2015). Given that residential stability is posited to increase familiarity, it is likely that it will have its strongest effect at a small scale.

Two other measures that may well exhibit micro-spatial effects are population density and vacant units. Both are important from a routine activities perspective. Vacant units are typically expected to provide opportunities for offenders to gather, as well as providing locations with a lack of guardians, hence increasing crime opportunities. Likewise, locations with low population density arguably have fewer eyes on the street (Jacobs 1961). The limited number of guardians in such locations would likely increase crime opportunities; however, this would arguably be more likely to occur at the micro scale of a block given that such guardians would be preventing specific crime events.

Nonetheless, the "small is better" approach and a focus exclusively on street blocks may not capture the entire story. Scale is particularly critical for racial/ethnic heterogeneity, inequality, and other distributional measures because they are dependent on comparisons between groups of people. These comparisons are implicitly tied to some spatial area. The set of people included in the calculations for these distributional measures is fundamentally dependent on the spatial scale of the process of interest. Using racial/ethnic heterogeneity as an example, the segregation that exists in many cities suggests that focusing into one area too narrowly might miss changes in the city landscape. Given that city blocks are often comprised mostly of the same ethnic/racial group, blocks will typically have strong within group homogeneity. To capture the racial/ethnic *heterogeneity* of an area necessarily implies a larger spatial area than these micro geographic units. The racial composition of the local block could potentially be much different from the surrounding environment. This suggests that street blocks and hot spots may be too small a geographic unit to capture the entire ecology of crime. Although it is uncertain how broad an area around the focal block will matter, it is nonetheless almost certain that what occurs on nearby street blocks will have important implications for the amount of crime on a particular street block. To the extent that these spatial processes get broad enough, it is even possible that simply measuring the characteristics at the broader meso scale of neighborhoods might not be broad enough to accurately capture the process.

*The broader environment*

It is worth emphasizing that at the same time that one burgeoning literature has focused on the importance of measuring crime at a very micro-scale, there is also growing awareness in neighborhood studies of crime that treating even meso-level neighborhoods as urban villages that have no interaction with surrounding neighborhoods is likely not theoretically appropriate. This viewpoint suggests that studying neighborhoods may even be too small a unit of analysis for understanding these processes. Nonetheless, it is worth emphasizing that for many studies the possibility of a broader spatial process outside of the neighborhood is simply treated as a methodological nuisance in which "spatial effects" must be accounted for when studying neighborhoods located in space. A more principled approach is to treat this as a theoretical issue, as one can specify the reasons why the characteristics of nearby neighborhoods might have consequences for the amount of crime in a focal neighborhood. Indeed, some scholars have suggested thinking of the linkages of neighborhoods to other neighborhoods in the city in terms of a social network perspective (Sampson 2004). In this perspective, there is a direct connection between the various neighborhoods of a city due to residential mobility, collective efficacy, and other processes. More recently, Hipp and Boessen (2013) used "egohoods" whereby neighborhood boundaries were overlapping and neighborhoods were spatially interdependent.

As suggested by Hunter's (1985) classic theoretical work, different forms of social control may have varying spatial dimensions. For example, parochial social control may be developed from friendship ties. While most research has focused on ties only within the neighborhood, more spatially distant ties have been shown to decrease neighborhood cohesion (Boessen et al., 2014). These more distant ties likely link different neighborhoods together to suggest a broader spatial scale, which may increase public social control. Institutional resources (i.e, schools, voluntary organizations, churches) and other forms of parochial social control suggest a broader spatial scale than the confines of one neighborhood because one neighborhood likely does not represent the total capacity for resource mobilization (Hunter, 1985; Janowitz, 1967). Neighborhoods that are entrenched with gangs and fights over turf imply processes of a broader spatial scale between neighborhoods (Brangtingham et al., 2012; Harding, 2010).[1]

Given the daily mobility patterns of people in urban areas, it seems extremely plausible that the characteristics of nearby neighborhoods can impact the amount of crime in a focal neighborhood. Just as offenders and victims do not constrain their activities to the block in which they reside, they also do not constrain their activities only to the meso-level neighborhood in which they live. For example, studies have shown that the daily activity patterns of residents (hence, potential targets or guardians) are spread relatively far geographically. Sastry, Pebley, and Zonta (2002) use the Los Angeles Families and Neighborhoods Study (LAFANS) to show that residents travel 1.37 miles on average to the grocery store and 8.15 miles to work. Similarly, in the National Household Travel Survey (U.S. Department of Transportation 2008) high school children travel an average of 6 miles to school. The evidence from the journey to crime literature makes clear that offenders do not simply commit crime within their own neighborhood, as studies suggest that offenders travel 2 miles on average to commit a crime (for a discussion of this research see Rossmo 2000). Even in neighborhoods that are walkable, it would be quite unlikely that the entire crime process contains residents only within the local area. Given that a typical census tract (what is often used as a proxy for a "neighborhood") is about 1.4 miles across, it is likely that much crime occurs outside an offender's own neighborhood. This suggests that a scale even larger than the particular neighborhood is likely important for understanding how much crime occurs in a neighborhood. Thus, there is likely a spatial process of crime, and it is not simply constrained to the notion of crime "diffusion" that has sometimes been posited in the literature (Cohen and Tita 1999).

[1] The broader spatial scale of these different processes suggests a particular spatial patterning that is nonrandom, but still arguably based on proximity. We leave these additional complications for future work.

*The role of land use in ecological models of crime*

Although much of the neighborhoods and crime literature focuses on the social characteristics of neighborhoods, the physical characteristics likely matter as well. This comes out of the defensible space literature (Newman 1972) and crime pattern theory (Brantingham and Brantingham 1984). In these perspectives, the physical characteristics of the area impact the location and timing of crime events. Crime pattern theory posits that potential offenders travel about the city and become aware of crime opportunities. For example, retail areas with many shoppers may provide more opportunities for crime, as would the nearby parking lots where they park their cars. The typical paths of offenders provide more opportunities for committing offenses, either because of the presence of more suitable targets, or fewer willing guardians.

Jacobs (1961) and New Urbanism approaches from the urban sociology and urban planning literatures suggest that mixed use neighborhoods are safer because of their walkability and their potential for more eyes on the street. This approach suggests that areas with both more commercial and residential land use should have less crime. On the other hand, Stark (1987) suggests that mixed use (i.e., residential and commercial land uses) nearby each other allows for more opportunities for deviance due to their walkability and the ease for young people to hang out, in tandem with classic social disorganization explanations including poverty, density, and rental units. In a similar vein, Sherman et al.'s (1989) hot spot paper suggests that crime is highly concentrated at a variety of places, but the vast majority appears to be at or near retail land use areas. These studies suggest a need to better understand different land use patterns.

Beyond the importance of measuring the physical characteristics of the local area is the question of the scale at which these characteristics might matter. One approach might suggest that land use features only operate at a very micro scale. For example, the presence of more industrial buildings may increase crime during off-peak hours because there are few guardians nearby to prevent crime events. A counter perspective is that whereas the land use of the local block matters, the land use in surrounding blocks, perhaps even to the meso level of the neighborhood, has important implications. In this view, the presence of industrial buildings on one block, may not have as deleterious effect if there are residential blocks nearby that provide many residents who are walking about in the area. In this case, the presence of persons from nearby areas could provide potential guardians to control the area. Or, the presence of residences nearby increases the potential targets for such industrial areas, since there would be more persons walking about. This would have an aggravating effect in which blocks with industrial buildings that have residential areas nearby would have more crime. Depending on the particular crime type, the people walking about may be the targets or the industrial establishments might be the targets. Regardless which of these processes is at work, this suggests a need to account for the land use composition of a larger geographic scale and to understand different crime types.

Interestingly, the neighborhoods and crime literature has paid limited attention to the importance of land use characteristics. Furthermore, the studies that have tested the effects of various land use patterns on neighborhood crime rates are often constrained to only testing the effect of a subset of possible land uses. Most common are the numerous studies that have shown that liquor stores and other alcohol outlets increase crime rates (Hipp 2007; Nielsen and Martinez 2003; Peterson, Krivo, and Harris 2000). These studies focus on how such outlets might affect the number and types of people who come to an area, as well as the possibly impaired state of these persons (if they have consumed alcohol), and the consequences for local crime rates. Another set of studies have looked at whether the neighborhood is in a central business district, generally finding that such neighborhoods have higher rates of crime (Bellair 2000; Warner and

Rountree 1997; Crutchfield 1989), although a study of Miami found no such effect (Nielsen and Martinez 2003). Occasional research has looked at the presence of retail outlets and suggested they increase violent crime and property crime (Lee et al., 2013)(see also Browning et al. 2010). Some studies have found that public housing is associated with more crime (Peterson, Krivo, and Harris 2000), whereas other studies have not found this effect (Lee et al., 2013). Other research has suggested that multifamily housing leads to more crime (Sherman et al., 1989)

One of the most exhaustive studies of land use and crime was that of Smith, Frazee and Davison (2000). This study focused on very small units of analysis—street blocks—and found that land use characteristics were associated with increased robberies. Although this study provided important insights of the effects of social disorganization theory and various land use characteristics at very small units of analysis, it did not take into account the characteristics of the larger area—either social demographic or physical characteristics. As a consequence, one goal of the present study is to examine these missing parts of the environment. It is still an empirical question whether it is only the characteristics of the local street block that matter. Indeed, Smith and Frazee found that the level of robberies on nearby street blocks affected the robbery rate on the focal street block, implying that spatial effects need to be examined.

*Summary*

Given the preceding discussion regarding the uncertainty of the appropriate scale for measuring ecological crime processes, we therefore test our models by including structural measures computed at three geographic units of analysis: blocks, block groups, and the spatial area within five miles around a block group. This allows us to simultaneously test for micro-processes (the block), meso-processes (the block group or tract), and wider area processes (the spatial area within five miles around the block group or tract). We tested models for both block groups and tracts, and generally found that the block group aggregations were more robust. We construct measures of the physical environment using five categories of land use, and aggregate these both to the micro unit of blocks as well as the meso unit of block groups.

**Data and Methods**

*Data*

Our study area is seven cities around the year 2000: 1) Chicago; 2) Cleveland; 3) Columbus; 4) Dallas; 5) Los Angeles; 6) San Francisco; 7) Tucson. The crime data were obtained directly from the police departments. These cities were not selected randomly, but rather are a convenience sample of cities, and as a result, this study does not generalize to the population of cities. The land use data were obtained from city and county planning, government, and assessor departments around the year 2000.[2] The cities and years of land use and crime data are presented in Appendix A. We included all blocks that were located in census tracts with nonzero population.

These cities vary along key dimensions. For example, San Francisco (195.3) and Chicago (146.5) have the greatest population density (in 100's per square mile) while Dallas (52.8) and Tucson (46.2) are the sparsest. Chicago and Cleveland have the most minority residents (Chicago has 39.8% black and 21.4% Latino) whereas San Francisco has the fewest (7.5% black and 13.3% Latino). Although Dallas has a mix of blacks and Latinos, Los Angeles and Tucson minorities are predominantly Latinos. Finally, Cleveland has more industrial land use than the other cities.

[2] The Chicago land use data that is used by county planning departments is from the Chicago Metropolitan Agency for Planning, and it is based on aerial photographs. All other land use data was in parcels: Cleveland (City Planning Commission); Columbus (Franklin County Assessor); Dallas (North Central Texas Council of Governments); Los Angeles (Southern California Association of Governments); San Francisco (City Planning Department); and Tucson (Pima County Assessor).

*Dependent Variables*

The dependent variables are from official police department data in each of the seven cities. Given that we have point data, we geocoded these events to latitude-longitude point locations, and then aggregated them to census blocks. We classified crime events into six crime types: aggravated assault, robbery, homicide, burglary, motor vehicle theft, and larceny.[3] We summed these measures over three years to minimize yearly fluctuations.

*Independent variables*

We aggregated the land use data into Census blocks and block groups, and appropriated (by area) the land use data to blocks when it was on a block boundary. We computed the proportion of the block or block group area classified into five land use measures: 1) residential[4]; 2) commercial; 3) industrial; 4) office space; 5) other (includes parking, parks, churches, open land, agriculture, hospitals, libraries, cemeteries, transportation, public buildings, etc.).[5] The five land use categories are exhaustive of all land uses, and we use "other" as the reference category in our models. While there are numerous land use categories across the 7 cities, we use these categories because of their consistency in measurement across the cities.

The socio-demographic characteristics come from 2000 U.S. Census data. We use data aggregated to blocks and block groups (or tracts). The measures constructed at both the block

[3] We do not include measures of sexual assault given the well-known reporting issues with such measures: consistently less than 40% of such incidents are reported to the police (Baumer and Lauritsen, 2010). San Francisco did not have homicide data available. For Chicago, we did not have aggravated assault data, but we only use general assault, which may or may not be aggravated.

[4] For most of these cities, we were able to distinguish between single- and multi-family units (except Chicago and San Francisco). The estimated coefficients were relatively similar for these two types of housing (although the single family housing unit variable typically was somewhat larger in absolute value than the multi-family unit variable) and that the coefficients for the other variables in the models were not substantively different when comparing a model including measures of the two types separately to a model with a combined measure. We therefore present the models with the combined measure, given that it allows us to include these two additional cities.

[5] It is not clear what represents a "mixed use" neighborhood: how much commercial must be present? How much residential? In other words, what is the spatial scale of mixed use? It is also unclear about combinations of different land uses: can a neighborhood be industrial and residential and still be mixed use? For our study, we chose not to include an explicit measure of "mixed use", but we include the various subtypes to more explicitly understand the driving land uses behind the "mixed use". Nonetheless, we see this as an intriguing challenge for future research.

and the block group level include: % vacant units, residential stability (combining standardized measures of % owners and % in the same housing unit 5 years previously), % African American, % Latino, population density (measured in 100's per square mile), and the % aged 16 to 29 (given that these are the prime ages of offenders). We also constructed a distributional measure of the racial/ethnic heterogeneity as a Herfindahl index of five racial/ethnic groupings (White, Black, Latino, Asian, and other races).

Some measures available from the U.S. Census, such as income, are not aggregated to blocks and are only available aggregated to block groups. We capture the economic environment of the block group with a measure of concentrated disadvantage. Using a principal components analysis, this measure combines: 1) % poverty; 2) % single parent households; 3) median household income; 4) median home value. We constructed a similar measure at the block level: % single parent households (the only measure available for blocks) along with imputed values of the other three measures.[6] We capture economic inequality (a distributional measure) with the Gini coefficient at the block group and tract aggregation.[7] As we describe below, we used the tract measure given that it consistently showed more robust relationships in the models. Given that the Census does not provide information on the income of households at the block level, it is not possible to measure inequality on blocks.

We calculated spatial lags of the socio-demographic variables by using a 5-mile inverse distance decay function. We also estimated spatial lags based on three other distance decay functions: 1) inverse distance decay capped at 2.5 miles; 2) biweight kernel distance decay

[6] We imputed these values using other information regarding the blocks, as well as information on the correlations among these variables at the block group level. This is the ecological inference approach, in which rather than simply using an areal apportioning approach, as is more common, it imputes the data to the smaller units based on the characteristics of those smaller units. For further discussion, please see the online supplemental information.
[7] We used the prln04.exe program provided by Francois Nielsen at http://www.unc.edu/~nielsen/data/data.htm, to account for the binned nature of the data.

capped at 5 miles; 3) biweight kernel distance decay capped at 2.5 miles (the biweight kernel can be represented as $(1-(dist(p,q)/r)^2)^2$ where r is the radius of the buffer and dist(p,q) is the distance in miles between the two blocks). The results were essentially identical using these various distance decay functions. For example, the correlations between the inverse distance decay measures at 2.5 versus 5 miles are all above .9. All of the spatial lags include information from neighborhoods that are outside the city boundary since estimates might be biased when only including information within cities (Wong 1997). We nonetheless included an indicator variable for blocks within 0.1 mile of the city border to control for this boundary issue.

The summary statistics for the variables used in the analyses are presented in Table 1. One feature to note is that the last four columns show 1) the intra-class correlation of the block measure (the degree of variance that is at higher geographic units than the block); 2) the correlation between the block and block group versions of the measure; 3) the correlation between the block and area surrounding the block group measures; 4) the correlation between the block group and area surrounding the block group measures. These values are conceptually interesting, as they give a sense of the degree to which these measures spatially vary across these geographic units. For example, the very high ICC value for percent black (.877) suggests that there is only minimal variability among the blocks within larger units for this measure, and the very high correlations of the block measure with the block group (.95) and surrounding area (.82) also show that there is only a modest amount of additional information contained in knowing the percent black in the block if one already knows the composition of the larger area. The measures of Latinos and concentrated disadvantage also exhibit relatively high spatial correlation, though not as high as for blacks. The measures of vacant units and population density exhibit the lowest spatial patterning of the socio-demographic measures, suggesting that there is less clustering of *people* in the environment than there is of the *types of people*.

<<<Table 1 about here>>>

**Methods**

Given that our outcome measures are the counts of crimes in blocks, we estimated multilevel Poisson models or multilevel negative binomial regression models when appropriate (using the menbreg command in Stata). In the initial models, the outcome variable is aggregated to the block group (or tract). For all models, we included the population within the unit as an exposure measure and this estimates the outcome as a crime rate. A model can be expressed as:

2.) $$E(y_{ij} \mid LU(k)_{ij}, x_{ij}) = LU(k)_{ij} B_1 + X_{ij} B_2 + X_j B_3 + WX_j B_4 + C B_5$$

where $y_{ij}$ is the expected crime count in the block, $LU(k)_{ij}$ is the proportion of the block that is composed of land use $k$ (of $K$-1 land use types of those defined earlier) with a $B_1$ vector of associations on the crime count, $X_{ij}$ is a vector of demographic measures of the block whose associations are captured in the $B_2$ vector, $X_j$ is a vector of demographic measures of the block group whose associations are captured in the $B_3$ vector, $WX_j$ is a vector capturing the demographic characteristics of the surrounding area whose associations are captured in the $B_4$ vector, C is a vector of the cities and their fixed effects are captured in the $B_5$ vector.

There was no evidence of collinearity problems in the models we estimated. Although some of the variance inflation factor values appear somewhat high (approaching 16), these high values are balanced by the very large sample size of blocks. As highlighted by O'Brien (2007), researchers are concerned about collinearity due to the possibility of inflated standard errors and hence unstable parameter estimates. Although some scholars only focus on variance inflation factor values (VIF) as a diagnostic for collinearity, VIF's are just one of four components that go into the standard error calculation: the other three are 1) the degree of variability in the variable of concern, 2) the sample size, and 3) the proportion of variance explained by the model. Given that we have a large sample size, we do not find problematic results when we use O'Brien's approach to calculate the degree to which the standard errors are inflated.[8] There was also no evidence of influential observations. We also tested our for spatial autocorrelation using Moran's I. We computed residuals by subtracting the predicted count for each unit from the actual count. The average Moran's I for all residuals averaged across cities was less than .04, which suggests very little substantive spatial autocorrelation remaining in the model after including the covariates.

We also estimated ancillary models in which we aggregated the neighborhood measures to tracts rather than block groups. These results were typically quite similar with the exception of our two distributional measures. First, racial/ethnic heterogeneity in the tract-only models was much stronger than in the block group-only models, suggesting that this measure better captured this construct (although the results were very similar using either block group or tract in the full models). Second, inequality was much stronger when measured at the tract level compared to the block group level in all models, suggesting this is a more appropriate aggregation. We therefore used the inequality measure aggregated to tracts in our final models.

We estimated two sets of models. The first set of models adopts the common approach in the neighborhoods and crime literature of aggregating crime to meso-level units (block groups, in this case) that are expected to approximate neighborhoods. We refer to these as our baseline models and they exclude the block-level measures. The second set of models utilize our full specification including micro-level demographic measures at the level of blocks, as well as

[8] For example, our largest VIF value was about 16. However, a variable with a VIF value of 2 in a sample of 200 would need to have a VIF of 735.3 in our sample of 73,010 blocks to equally impact the imprecision of the estimates. Thus, sample size quite dramatically improves the precision of estimates (Goldberger 1991).

accounting for the land uses of the block and block group. In instances in which there were considerable differences between the baseline model and our full model specifications, we estimated ancillary models to assess why such changes are observed. Specifically, we estimated intervening models which: a) added block level demographic measures but not the land use measures; and b) added block level land use measures but not the block demographic measures. To avoid an abundance of tables, we only describe the pattern of these intervening results when they account for why differences occurred, and do not present them in tables.

**Results**

We discuss the results for each of our covariates in turn, first viewing the results for the baseline models specified at the block group level (models 1a, 2a, and 3a in Table 2), and then discussing the spatial pattern in the full models specified at the block level (models 1b, 2b, and 3b in Table 2). Three of the structural measures in the model show distinctly micro relationships: residential stability, percent vacant units, and population density. For example, whereas a higher percentage of vacant units in the block group is associated with more aggravated assaults in the block group models ($\beta$ = 1.048 in models 1a in Table 2), the size of this association when including the block level measures is reduced 40% in the aggravated assault model ($\beta$ = .604 in model 1b). Instead, it is the presence of more vacant units on the block that is associated with higher aggravated assaults, robberies and homicides. Likewise in the burglary model the significant block group relationship for percent vacant units in model 1a in Table 3 is 1/3 the size when including the block measures in model 1b. The presence of more vacant units on the block is associated with higher levels of all three property crimes. Whereas the presence of more vacant units in surrounding areas appeared associated with more violence in the initial models (models 1a, 2a and 3a), the broader spatial pattern is not present for robbery in the full models.

<<<Table 2 about here>>>

<<<Table 3 about here>>>

Residential stability also exhibits a very micro relationship with crime, and it is associated with less crime. Although it appears in the baseline models that block groups with more stability have fewer robberies, we see that this is in fact a micro association in which blocks with more stability have fewer aggravated assaults and robberies (models 1b and 2b in Table 2). In the property crime models, stability again exhibited a micro-level association: blocks with more stability had lower rates of all three property crimes (Table 3). There was some evidence that more stability in the broader block group was associated with fewer motor vehicle thefts, but this relationship was not detected for any of the other crimes.

The pattern for population density was very distinct. More population density decreases all six crime types in the block group models. However, when including the block level measures the associations at the block group level actually change sign in the violent crime models, and become nonsignificant in the burglary and larceny models. Whereas more population density on the block is associated with lower levels of all crime types, more population density in the block group is associated with more violent crimes. Greater population density in the surrounding area is also associated with more crime for all of the crime types except motor theft.

Not all of the variables in these models exhibit such pronounced micro-scale associations. For example, racial minorities in block groups are associated with increases in violent crime rates and this finding remains significant in our full models, although often weaker.

Furthermore, the micro associations for racial minorities are much weaker: although blocks with more Latinos or African Americans have higher aggravated assault rates, they do not have significantly more robberies and homicides. In the property crime models, the results differ over crime type. The presence of more African Americans in the block, the block group, and the surrounding area are all associated with more motor vehicle thefts, but there are no such relationships for the other property crimes. And whereas it appears in the block group models that Latinos are associated with lower burglary and larceny rates, this is actually mostly a micro-scale relationship as blocks with more Latinos have less crime, and the presence of Latinos in the block group only remains significant for larcenies.

When we estimated intervening models to assess why these changes occur in our model specification, we found that it is the inclusion of the demographic measures at the block level that alters these results. Thus, including land use measures aggregated to block groups do not change the racial composition results at all. Furthermore, including land use measures aggregated to the micro unit of blocks only modestly changes these results; it is only when including the demographic characteristics measured at the more micro scale that these racial composition block group relationships change.

The spatial patterning for racial/ethnic heterogeneity is particularly distinct. On the one hand, in baseline models in which we aggregated all of our measures to *tracts*, we obtained the common finding in the literature that tracts with more racial/ethnic heterogeneity have higher rates of all six of these crime types (not shown). On the other hand, in preliminary full models we found not only block group racial/ethnic heterogeneity associated with more crime, but also that *blocks* with more heterogeneity actually have *lower* crime rates. This implies that the highest crime rate combination occurs on a highly segregated block (that is, a homogenous racial composition) within a block group with a high level of racial/ethnic heterogeneity. Thus, it is this micro-clustering by race within the meso-context of heterogeneity that results in the most violence. To further explore this relationship, we also included an interaction between block and block group racial/ethnic heterogeneity in our full models. We visually plot this result for the aggravated assault model in Figure 1, and as shown, whereas the lowest assault rate occurs in a block group with low heterogeneity, regardless of the heterogeneity of the block (the left side of the figure), the highest assault rate occurs on a homogeneous block contained within a heterogeneous block group (the right side of the figure). Although not shown, the pattern was similar for all other crime types except homicide. We emphasize that the appropriate interpretation is that heterogeneity increases due to composition changes of whites, Asians, or other race (holding constant the percent black and Latino) are associated with higher crime rates.

<<<Figure 1 about here>>>

The results for concentrated disadvantage also exhibit a spatially diffuse association. Block groups with higher levels of concentrated disadvantage have higher violent crime rates in our baseline models, and these relationships remain robust when including the micro-level measures. We also see that blocks with higher levels of concentrated disadvantage in the surrounding area have higher aggravated assault and robbery rates. In the full model there is a micro association in which blocks with more concentrated disadvantage have higher aggravated assault rates beyond the block group association. However, there is no evidence of a micro-spatial pattern for concentrated disadvantage for the other two violent crimes, and blocks with higher rates of concentrated disadvantage actually have *lower* property crime rates than other blocks. Instead, it is the broader spatial pattern that matters for property crimes: higher concentrated disadvantage in the block group is associated with higher burglary and motor vehicle theft rates, and higher concentrated disadvantage in the surrounding areas is associated with more larcenies.

We find that inequality has a relatively strong relationship with these crime types. Tracts with more income inequality have more aggravated assaults, robberies, burglaries and larcenies. Furthermore, inequality was always stronger in our full models than in the baseline models. When examining the results from the intervening models, we determined that it was the introduction of the land use measures at the micro scale that brought about this change in the inequality measure, resulting in findings more in line with theoretical expectations.

Finally, we find that the land use measures generally show micro associations, although there are still some spatially diffuse relationships in the models. Whereas the measures aggregated to block groups appear to be robust in the baseline models, these results typically weaken, and sometimes reverse, in the full models. Blocks with more commercial buildings have more of all six crime types; the presence of commercial buildings in the broader block group suggests additional higher levels of aggravated assaults, robberies, and motor vehicle thefts. A block with all commercial buildings has about twice as many aggravated assaults as a block with none, 175% more homicides, about 9 times more robberies, and between 250% and 400% more property crimes. The presence of office buildings on the local block has a stronger impact for increases in property crimes than on violent crimes (only significant for robberies). The presence of office buildings in the broader block group is associated with more burglaries, larcenies, and aggravated assaults. Interestingly, the relationship between industrial land use and more violence is entirely contained in the block group measure (the block measure is nonsignificant or associated with decreases in crime). However, industrial land use in blocks does appear to act as an attractor for property crimes, and burglaries and motor vehicle theft rates are even higher if the surrounding block group also has industrial usage. Residential units are associated with lower levels of these crime types at the relatively micro scale of blocks. A block with all residential units will have about 80 percent fewer aggravated assaults and robberies than a block with none. And whereas it appeared in our baseline models that block groups with more residential units have lower motor vehicle theft and larceny rates, this is in fact entirely micro as this block group coefficient is not associated with less crime in the full models. In fact, in the full models the presence of more residential land use in the block group is associated with more crime for four of the crime types, controlling for the level of residential units in the block.

Table 4 provides a summarized version of the results of the full models by noting whether the relation of a construct with a crime type significantly increases (“+”), significantly deceases (“-”), or neither. For the measures of percent black and Latino, we see that they are relatively consistently associated with higher crime at the block group level, but weak or mixed results at the block level. And for both, increases in their spatial lags suggest more property crimes and robberies (acquisitive crimes). Racial/ethnic heterogeneity, holding constant the percent black or Latino, is quite robust with the interaction between the block and block group displayed earlier in Figure 1. We also see that concentrated disadvantage is robustly associated with more crime when measured at the neighborhood or broader area, but the block measure has a weaker association for violent crime. Block disadvantage is also associated with less property crime. In contrast, the residential stability measure is robustly associated with less crime at the block level, but weak and mixed results for the larger aggregations. Vacant units are also micro, as the block measure is always associated with higher crime rates. The neighborhood and broader area measures of vacant units show occasional significant increases in these crime types. The population density measure also shows a pronounced spatial pattern: higher density in the block is associated with lower crime rates, but higher density in the neighborhood and broader area is often related to higher crime rates. The one notable exception is motor vehicle thefts: the presence of more population density *or* vacant units in the neighborhood and broader area is associated with *fewer* motor vehicle thefts, which is a strikingly different pattern than for the other crime types. For the land use measures, commercial land use shows a robust relationship at the block level to suggest more crime, but a weaker relationship at the neighborhood level, although still suggesting more crime. Industrial land use is less micro: whereas industrial land use in the block increases *property* crimes, the neighborhood composition is associated with higher rates of all crimes. The presence of office land use, especially at the block level, is associated with higher rates of acquisitive crimes (property crimes and robberies). Finally, residential land use blocks have lower rates of all crime types; however, there is no additional contextual association, as the proportion residential in the neighborhood is associated with marginally more of several crime types.

**Conclusion**

In this study we have argued that scholars of the ecology of crime need to take much more seriously the level of aggregation when considering the association between neighborhood structural characteristics and rates of crime. Whereas this suggestion is not entirely new (Hipp 2007, Taylor 2015), few studies have rigorously tested this. We showed that considerable insights are provided by taking into account the social characteristics not only at the *micro-scale* of the local block, but also simultaneously accounting for the *meso scale* of the neighborhood as well as the *broader area surrounding* a neighborhood. The pattern of results was altered—sometimes quite considerably—by accounting for these varying scales. When we included the characteristics of the physical environment via land use, they strongly predicted the location of crime, and for some measures they altered the observed relationships of the socio-demographic measures.

Racial/ethnic heterogeneity showed particularly striking pattern of results. On the one hand, a model adopting the more traditional approach measuring heterogeneity in the meso area of the neighborhood and the broader spatial area around the neighborhood found that heterogeneity increased neighborhood crime (consistent with the literature). Yet, when taking into account the micro spatial area we found an interesting pattern in which heterogeneity at the block level decreased crime, whereas heterogeneity in the meso level of the neighborhood increased it. In high heterogeneity block groups, it is the homogeneous blocks within them that will have the most crime. This suggests that segregation *within* a neighborhood has a particularly strong association with crime, a finding that has not been documented in the literature. Thus, block homogeneity, surrounded by high levels of heterogeneity, leads to the highest rates of crime.

This racial homogeneity finding suggests a particular form of both social and spatial isolation (Wilson 1987). Indeed, research often suggests economic isolation from the rest of the *city*, but in this case, we found a pattern within the same neighborhood. Additionally, while we might expect most ties will exist on the block (Taylor 1997), research has suggested that ties outside of the neighborhood may be particularly salient for garnering resources (Bellair 1997). This finding suggests that some homogenous neighborhoods may be at a particular disadvantage if the surrounding area is heterogeneous. Thus, for these residents, there is likely more social distance (and fewer ties) between the meso area and the homogenous block. As such, our findings suggest the importance for future research to measure the spatial distribution of ties. Strong within group preferences on the local block may also create a tension with the surrounding area, thereby residents from outside the block are more likely to travel to commit crimes in this area. Future research might examining gang territories since these areas may be homogenous areas within high crime heterogeneous areas (see also Pattillo 1998).

Another key finding was that some of these characteristics appear to operate at the much smaller geographic scale of blocks. For example, it is the presence of more residential stability on the local block that sharply reduces the level of crime on the block. When accounting for this micro-scale, the meso association of stability in the block group weakened considerably to suggest a protective process at a very micro scale. Residential stability may create an environment that provides more ties among neighbors and therefore familiarity; these micro findings for stability are consistent with our earlier discussion that familiarity is more likely to be generated at the micro block level. Indeed, the ability to provide eyes on the street is likely a micro spatial process, and might explain why population density and vacant units both exhibited micro associations.

A striking finding was the complicated relationship between population density and property crime. Although higher levels of population density at the micro scale of the block reduced all types of crime, greater population density in the neighborhood and surrounding area was actually associated with *higher* violent crime rates. Population density in the broader area was associated with increases in burglary and larceny, as well. The complexity surrounding this relationship suggests a nonlinear scaling of population density, and therefore there likely exists enormous heterogeneity in guardianship across the city (Butts et al., 2012).

Some of the strongest predictors in our models were from our land use measures. Although land use appears to have a limited presence in neighborhoods literature, this study helps to reinforce the importance of physical characteristics, and the potential of incorporating this more precise spatial information into our models.  Not only did we find strong associations for the land use measures, but they also sharpened some of our estimates.  Accounting for the land use is important for specifying these models.  This is somewhat unsurprising, given that the land use variables generally had quite strong associations.  Thus, blocks with more residential units had much *lower* rates of all crime types, blocks with more commercial land use had *higher* rates of crime, and blocks with more office and industrial land use generally had higher property crime rates.

It is surprising that land use measures have gained little traction in prior research since they likely regulate many of the social processes that are expected to control crime.  Hipp (2007) demonstrated that different neighborhood processes have different associations depending on the unit employed by researchers, and he suggests using theory to guide the geographic scale of neighborhood processes.  Extending this line thought, Butts et al., (2012) and Hipp et al., (2013) suggest that social networks and population are unevenly spread over different neighborhoods in the city: in one paper they found that simulated networks had consequences for actual crime rates.  While the main goal of those two papers was simulation of social networks, the current study extends this area of work by incorporating physical aspects of neighborhoods—land uses—for understanding neighborhood processes. Most often neighborhood researchers have relied exclusively on Census data or large-scale survey data typically paired with Census data. The incorporation of land use data is an approach that can be adopted to understand the physical aspects of neighborhoods that help to potentially situate many social processes.  An area for future work is to make more theoretical progress on the how physical characteristics work in tandem with the social characteristics of the area to create criminal opportunities.  Future research might unpack the land uses types to better understand the diversity of uses over the city.

While this paper has exclusively focused on *spatial* scale, we have little understanding the *temporal* scale of different social processes (Abbott, 2001, Taylor 2015). Given that some research has focused on the situational aspects of offending (e.g., see Bernasco et al. 2013), future research might also examine the situational aspects of when various land uses are risky throughout the day. A challenge for this area of research is to incorporate explanations (i.e. "causes") for crime and other social phenomenon that are not exclusively within one discrete concurrent spatial-temporal context and more explicitly incorporate differing spatial temporal scales.

Whereas social disorganization theory might suggest a *long-term macro* socio demographic neighborhood change approach, routine activities and problem oriented policing work has mostly examined *short-term micro* processes on street segments or in crime hot spots. The findings from this study and many others suggest that both approaches have merit, but our study suggests that a focus exclusively on micro or macro does not tell the entire story for the social ecology of crime. In fact, zooming in to one area appears problematic because of the geographic scale of different neighborhood processes. Future research might more explicitly examine the micro/macro and short/long -term social dynamics of neighborhood processes.

A question arises regarding how robust our results were across cities. We assessed this by estimating separate models for each city and found relatively consistent results. Although not shown, almost none of the results in our Tables were driven by just one or two cities: the one exception was the association of population density and motor vehicle theft, which was only found in Los Angeles. The significant results we found were typically similar across cities, with the exception the spatial lags of vacant units and population density for motor vehicle theft, which were quite mixed across cities. The spatial patterns for motor vehicle theft were somewhat unique from the results for the other crime types, suggesting an avenue for future research of why this might be the case. The different results for Los Angeles—a city whose spatial scale is notably defined by sprawl and a car culture—suggests that future work exploring these spatial processes within very different macro settings may be useful.

We acknowledge that this study does have some limitations. First, the observed differences shown here may not necessarily be due just to scaling issues. The findings here assume all variables scale linearly. Second, another assumption is that the boundaries are conceptually meaningful for understanding how to differentiate between units. Indeed, the boundaries between units might be endogenous (Rey et al. 2011). Third, we were constrained to using official reports of crime incidents, and it is well-known that there is under-reporting of such data (Lynch and Addington 2007; MacDonald 2001). Nonetheless, there is evidence that such under-reporting for Type 1 crimes is not systematically related to the characteristics of neighborhoods, suggesting that the coefficients may be relatively unbiased (Baumer 2002). Fourth, similar to almost all neighborhoods and crime studies the crime inducing and reducing processes suggested by different land uses were not explicitly observed, and future research will want to more explicitly capture these processes. Future research might incorporate longitudinal data, as well as a more explicit understanding for how crime may impact changes in land use patterns over the long term. Finally, we only studied these ecological processes in seven cities, and future research will need to test these models on additional cities.

Although the importance of scale has been known for decades, this study reinforces the consequences of it for studies of the ecology of crime to suggest a mutliscale approach. Using data from seven cities, we routinely observed different associations when rescaling our measures, and thus we would have reached different conclusions if we had only focused on the micro or the meso. We also found that land use characteristics seem to be crucially important for understanding the spatial distribution of crime. The findings emphasize the point that there is no reason to suspect that all neighborhood processes necessarily operate on the same scale, and the determination of the proper scale for a measure should be dependent upon theory and the research question of interest (Hipp 2007). The challenge for future research will be to balance the interdependence between micro and macro neighborhood processes.

**References:**

Abbott, A. (2001). Time Matters: On Theory and Method. University of Chicago Press.

Baumer, Eric P. 2002. Neighborhood Disadvantage and Police Notification by Victims of Violence. *Criminology* 40 (3):579-616.

Baumer, Eric and Janet L. Lauritsen. 2010. "Reporting Crime to the Police, 1973–2005: A Multivariate Analysis of Long-Term Trends in the National Crime Survey (NCS) and National Crime Victimization Survey (NCVS)." *Criminology* 48:131-185.

Bellair, Paul E. 1997. Social Interaction and Community Crime: Examining the Importance of Neighbor Networks. *Criminology* 35 (4):677-703.

———. 2000. Informal Surveillance and Street Crime: A Complex Relationship. *Criminology* 38 (1):137-167.

Bernasco, Wim, and Richard L. Block. 2009. Where Offenders Choose to Attack: A Discrete Choice Model of Robberies in Chicago. *Criminology* 47 (1):93-130.

———. 2011. Robberies in Chicago: A Block-Level Analysis of the Influence of Crime Generators, Crime Attractors, and Offender Anchor Points. *Journal of Research in Crime and Delinquency* 48 (1):33-57.

Bernasco, Wim, S. Ruiter, G. J. N. Bruinsma, L. J. R. Pauwels, and F. M. Weerman. 2013. Situational Causes of Offending: A Fixed Effects Analysis of Space-Time Budget Data. *Criminology* 51 (4).

Boessen, Adam, and Elizabeth Cauffman. Moving from the neighborhood to the cellblock: The impact of youth's neighborhoods on prison misconduct. *Crime & Delinquency*, In Press. doi: 10.1177/0011128713478131

Boessen, Adam., John R. Hipp, Emma J. Smith, Carter T. Butts, Nicholas N. Nagle, and Zack W. Almquist. 2014. Networks, space, and residents' perception of cohesion. *American Journal of Community Psychology*, 53:447–461.

Brantingham, Paul J., and Patricia L. Brantingham. 1984. *Patterns in Crime*. New York: MacMillan.

Brantingham, P. Jeffery, George E. Tita, Martin B. Short, and Shannon Reid. 2012. The ecology of gang territorial boundaries. *Criminology*, 50(3):851–885.

Browning, Christopher R., Reginald A. Byron, Catherine A. Calder, Lauren J. Krivo, Mei-Po Kwan, Jae Yong Lee, and Ruth D. Peterson. 2010. Commercial Density, Residential Concentration, and Crime: Land Use Patterns and Violence in Neighborhood Context. *Journal of Research in Crime and Delinquency* Forthcoming.

Bursik, Robert J. 1988. Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects. *Criminology* 26 (4):519-551.

Butts, Carter T., Ryan M. Acton, John R. Hipp, and Nicholas N. Nagle. 2012. Geographical Variability and Network Structure. *Social Networks* 34 (1):82-100.

Cohen, Jacqueline, and George Tita. 1999. Diffusion in Homicide: Exploring a General Method for Detecting Spatial Diffusion Processes. *Journal of Quantitative Criminology* 15 (4):451-493.

Crutchfield, Robert D. 1989. Labor Stratification and Violent Crime. *Social Forces* 68 (2):489-512.

Goldberger, Arthur S. 1991. *A course in econometrics*. Cambridge MA: Harvard.

Grannis, Rick. 2009. *From the Ground Up: Translating Geography into Community through Neighbor Networks*. Princeton: Princeton.

Harding, David .J. 2010. Living the Drama: Community, Conflict, and Culture among Inner-City Boys. University of Chicago Press, Chicago, IL.

Hipp, John R. 2007. Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point. *American Sociological Review* 72 (5):659-680.

———. 2007. Income Inequality, Race, and Place: Does the Distribution of Race and Class within Neighborhoods affect Crime Rates? *Criminology* 45 (3):665-697.

Hipp, John R. and Adam Boessen. 2015. "Neighborhoods, Networks, and Crime." in Challenging Criminological Theory: The Legacy of Ruth Kornhauser, edited by F. T. Cullen, P. Wilcox, R. J. Sampson, and B. Dooley: Transaction.

Hipp, John R., Carter T. Butts, Ryan M. Acton, Nicholas N. Nagle, and Adam Boessen. (2013). Extrapolative simulation of neighbor- hood networks based on population spatial distribution: Do they predict crime? Social Networks. 35: 614-625.

Hunter, Albert. 1985. Private, parochial, and public social orders: the problem of crime and incivility in urban communities. In Suttles, G. and Zald, M., editors, The challenge of social control: citizenship and institution building in modern society, pages 230–242. Norwood, NJ: Ablex.

Jacobs, Jane. (1961). *The Death and Life of Great American Cities*. Vintage Series. New York: Vintage Books.

Janowitz, Mark. *The Community Press in an Urban Setting*. Chicago: University of Chicago Press, 1967.

Krivo, Lauren J., and Ruth D. Peterson. 1996. Extremely Disadvantaged Neighborhoods and Urban Crime. *Social Forces* 75 (2):619-648.

Lee, Barrett A., Sean F. Reardon, Glenn Firebaugh, Chad R. Farrell, Stephen A. Matthews, and David O'Sullivan. 2008. Beyond the census tract: Patterns and determinants of racial segregation at multiple geo- graphic scales. *American Sociological Review* 73:766–91.

Lynch, James P., and Lynn A. Addington, eds. 2007. *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR*. Cambridge: Cambridge University Press.

MacDonald, Ziggy. 2001. Revisiting the Dark Figure. *British Journal of Criminology* 41 (1):127-149.

McCue, Kenneth F. 2011. "Creating California's Official Redistricting Database." http://statewidedatabase.org/d10/Creating%20CA%20Official%20Redistricting%20Database.pdf

Mears, Daniel P. and Avinash S. Bhat. 2006. No community is an island: The effects of resource deprivation on urban violence in spatially and socially proximate communities. Criminology, 44(3):509– 548.

Newman, Oscar. 1972. *Defensible Space: People and Design in the Violent City*. London: Architectural Press.

Nielsen, Amie L., and Ramiro Martinez. 2003. Reassessing the Alcohol-Violence Linkage: Results from a Multiethnic City. *Justice Quarterly* 20 (3):445-469.

O'Brien, Robert M. 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity* 41 (5):673-690.

Oberwittler, D., and Per-Olof H. Wikstrom. 2009. Why small is better. Advancing the study of the role of behavioral contexts in crime causation. In *Putting crime in its place. Units of analysis in geographic criminology*, edited by D. Weisburd, W. Bernasco and G. Bruinsma. New York: Springer.

Pattillo, Mary E. 1998. Sweet mothers and gangbangers: Managing crime in a black middle-class neighborhood. *Social Forces* 76 (3):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 (1):31-63.

Rey, Sergio J., Luc Anselin, David C. Folch, Daniel Arribas-Bel, Myrna L. Sastré Gutiérrez, and Lindsey Interlante. 2011. Measuring Spatial Dynamics in Metropolitan Areas. *Economic Development Quarterly* 25 (1):54-64.

Rossmo, D. Kim. 2000. *Geographic Profiling*. Boca Raton: CRC Press.

Sampson, Robert J. 2004. Networks and Neighbourhoods: The Implications of Connectivity for Thinking About Crime in the Modern City. In *Network Logic: Who Governs in an Interconnected World?*, edited by H. McCarthy, P. Miller and P. Skidmore. London: Demos.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science* 277 (August 15):918-924.

Sastry, Narayan, Anne R. Pebley, and Michela Zonta. 2002. *Neighborhood Definitions and the Spatial Dimension of Daily Life in Los Angeles.* Santa Monica, CA: RAND.

Slocum, Lee A., Andres F. Rengifo, Tiffany Choi, and Christopher R. Herrmann. .2013.. The elusive relationship between community organizations and crime: an assessment across disadvantaged areas of the South Bronx. Criminology, 51(1):167–216.

Sherman, Lawrence W., Partick R. Gartin, and Michael E. Buerger. 1989. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–55.

Smith, William R., Sharon Glave Frazee, and Elizabeth L. Davison. 2000. Furthering the Integration of Routine Activity and Social Disorganization Theories: Small Units of Analysis and the Study of Street Robbery as a Diffusion Process. *Criminology* 38 (2):489-523.

Stark, Ropdney. (1987). Deviant places: A theory of the ecology of crime. *Criminology*, 25(4):893–909.

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 (1):113-155.

Taylor, Ralph B. 2015. Community Criminology: Fundamentals of Spatial and Temporal Scaling, Ecological Indicators, and Selectivity Bias. New York: New York University Press.

U.S. Department of Transportation. 2008. NHTS Brief: National Household Travel Survey: Federal Highway Administration.

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 (4):520-536.

Weisburd, David, Elizabeth Groff, and Sue-Ming Yang. 2012. *The Criminology of Place*. New York, Oxford.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Wong, David W. S. 1997. Spatial Dependency of Segregation Indices. *Canadian Geographer* 41 (2):128-136.

## Tables and Figures

Table 1. Summary statistics for measures used in analyses

| | Block | | Block group | | Surrounding area | | ICC | Correlations | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | | Block and BG | Block and nearby | BG and nearby |
| ***Dependent variables*** | | | | | | | | | | |
| Aggravated assault | 2.99 | 8.15 | | | | | 0.286 | | | |
| Robbery | 1.65 | 4.19 | | | | | 0.245 | | | |
| Homicide | 0.06 | 0.27 | | | | | 0.287 | | | |
| Burglary | 3.48 | 6.19 | | | | | 0.238 | | | |
| Motor vehicle theft | 3.15 | 6.37 | | | | | 0.231 | | | |
| Larceny | 9.00 | 27.81 | | | | | 0.283 | | | |
| ***Independent variables*** | | | | | | | | | | |
| Percent black | 23.0% | 35.1% | 23.1% | 33.7% | 21.7% | 22.7% | 0.877 | 0.947 | 0.824 | 0.853 |
| Percent Latino | 25.0% | 30.2% | 26.1% | 28.7% | 29.3% | 20.0% | 0.753 | 0.889 | 0.726 | 0.783 |
| Racial/ethnic heterogeneity | 32.84 | 21.92 | 37.68 | 20.35 | 55.81 | 11.99 | 0.544 | 0.729 | 0.442 | 0.514 |
| Concentrated disadvantage | 0.01 | 0.13 | -0.07 | 1.00 | -0.05 | 1.01 | 0.628 | 0.785 | 0.609 | 0.731 |
| Residential stability | 0.00 | 0.69 | 0.00 | 0.74 | 0.00 | 0.66 | 0.559 | 0.718 | 0.489 | 0.603 |
| Percent vacant units | 6.1% | 9.4% | 6.3% | 6.4% | 6.1% | 2.6% | 0.222 | 0.480 | 0.353 | 0.544 |
| Population density | 102.0 | 176.0 | 121.6 | 111.9 | 84.8 | 51.5 | 0.207 | 0.441 | 0.311 | 0.664 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Percent aged 16 to 29 | 21.2% | 12.3% | 22.1% | 9.2% | | | 0.361 | 0.555 |
| Economic inequality | | | 0.87 | 0.14 | | | | |
| ***Land use features*** | | | | | | | | |
| Percent commercial buildings | 4.7% | 14.4% | 5.5% | 8.6% | | | 0.185 | 0.427 |
| Percent industrial buildings | 3.0% | 13.7% | 5.2% | 13.3% | | | 0.395 | 0.515 |
| Percent office buildings | 0.8% | 6.3% | 1.1% | 4.4% | | | 0.363 | 0.460 |
| Percent residential | 74.8% | 30.9% | 64.7% | 25.5% | | | 0.353 | 0.481 |

*N = 73,010 blocks across seven cities (Chicago; Cleveland; Columbus; Dallas; Los Angeles; San Francisco; Tucson).*
Note: Crime rates are logged.

**Table 2. Multilevel negative binomial regression models for three types of violent crime; including block, block group, and spatially lagged measures of demographics, and block and block group measures of land use variables**

| | (1a) | | (1b) | | (2a) | | (2b) | | (3a) | | (3b) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Assault | | Assault | | Robbery | | Robbery | | Homicide | | Homicide | |
| ***Block measures*** | | | | | | | | | | | | |
| Percent black | | | 0.8932 | ** | | | 0.1736 | † | | | 0.5277 | † |
| | | | (10.01) | | | | (1.76) | | | | (1.85) | |
| Percent Latino | | | 0.3770 | ** | | | -0.0796 | | | | 0.2577 | |
| | | | (4.96) | | | | -(0.90) | | | | (1.01) | |
| Racial/ethnic heterogeneity | | | 0.3857 | ** | | | 0.1269 | | | | -0.0925 | |
| | | | (3.64) | | | | (1.00) | | | | -(0.31) | |
| Concentrated disadvantage | | | 0.5443 | ** | | | 0.1789 | | | | 0.4675 | |
| | | | (4.55) | | | | (1.33) | | | | (1.49) | |
| Residential stability | | | -0.1533 | ** | | | -0.1071 | ** | | | -0.0654 | |
| | | | -(7.96) | | | | -(4.92) | | | | -(1.19) | |
| Percent vacant units | | | 1.0294 | ** | | | 1.0152 | ** | | | 1.4572 | ** |
| | | | (9.32) | | | | (8.91) | | | | (5.90) | |
| Population density | | | -1.4618 | ** | | | -1.8888 | ** | | | -1.9769 | ** |
| | | | -(10.35) | | | | -(12.51) | | | | -(8.24) | |
| Percent aged 16 to 29 | | | -0.0582 | | | | 0.0747 | | | | 0.6808 | † |
| | | | -(0.47) | | | | (0.58) | | | | (1.94) | |
| Proportion commercial buildings | | | 0.7765 | ** | | | 2.3459 | ** | | | 1.0137 | ** |
| | | | (10.65) | | | | (28.82) | | | | (5.82) | |
| Proportion industrial buildings | | | -0.2247 | * | | | 0.0216 | | | | 0.0819 | |
| | | | -(2.56) | | | | (0.21) | | | | (0.40) | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proportion office buildings | | | 0.2523 | | | | 0.7069 | ** | | | -0.3381 | |
| | | | (1.53) | | | | (3.20) | | | | -(0.71) | |
| Proportion residential | | | -1.4742 | ** | | | -1.9153 | ** | | | -0.7646 | ** |
| | | | -(31.43) | | | | -(39.07) | | | | -(7.45) | |
| Case is within 0.1 mile of city border | | | 1.7287 | ** | | | 0.7226 | | | | -2.6289 | |
| | | | (2.88) | | | | (0.84) | | | | -(1.46) | |
| ***Block group measures*** | | | | | | | | | | | | |
| Percent black | 1.1365 | ** | 0.5170 | ** | 0.7410 | ** | 0.6867 | ** | 2.3011 | ** | 1.8658 | ** |
| | (20.94) | | (5.03) | | (10.30) | | (5.75) | | (12.36) | | (5.70) | |
| Percent Latino | 0.7084 | ** | 0.4114 | ** | 0.2619 | ** | 0.3355 | ** | 1.4770 | ** | 1.3457 | ** |
| | (13.67) | | (4.73) | | (3.82) | | (3.07) | | (8.85) | | (4.75) | |
| Racial/ethnic heterogeneity | 0.1350 | ** | 0.9466 | ** | 0.1726 | ** | 0.8183 | ** | 0.3369 | * | 0.4990 | † |
| | (2.88) | | (9.37) | | (2.77) | | (6.67) | | (2.44) | | (1.92) | |
| Concentrated disadvantage | 0.2695 | ** | 0.2475 | ** | 0.1314 | ** | 0.1456 | ** | 0.2715 | ** | 0.2938 | ** |
| | (16.97) | | (11.84) | | (6.31) | | (5.52) | | (5.82) | | (5.77) | |
| Residential stability | -0.0469 | ** | 0.0925 | ** | -0.0856 | ** | -0.0204 | | 0.0495 | | 0.1448 | * |
| | -(2.75) | | (3.92) | | -(3.84) | | -(0.69) | | (1.00) | | (2.33) | |
| Percent vacant units | 1.4084 | ** | 0.6043 | ** | 0.5146 | ** | 0.0845 | | 2.0460 | ** | 0.5723 | |
| | (9.26) | | (3.30) | | (2.58) | | (0.40) | | (5.84) | | (1.64) | |
| Population density | -1.0758 | ** | 0.4354 | ** | -1.3667 | ** | 0.4510 | ** | -0.8308 | ** | 0.6301 | * |
| | -(13.82) | | (3.32) | | -(12.94) | | (2.96) | | -(3.26) | | (2.29) | |
| Percent aged 16 to 29 | -0.1580 | | -0.0596 | | 0.2060 | | 0.2351 | | -0.0841 | | -0.4491 | |
| | -(1.44) | | -(0.41) | | (1.48) | | (1.41) | | -(0.25) | | -(1.16) | |
| Economic inequality | 0.1878 | ** | 0.4040 | ** | 0.3131 | ** | 0.7378 | ** | 0.0282 | | 0.0975 | |
| | (3.12) | | (4.28) | | (3.93) | | (6.05) | | (0.16) | | (0.37) | |
| Proportion commercial buildings | 1.3943 | ** | 0.3914 | ** | 3.6855 | ** | 0.7940 | ** | 0.6476 | * | -0.4055 | |
| | (13.74) | | (3.66) | | (25.29) | | (5.07) | | (2.39) | | -(1.29) | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proportion industrial buildings | 0.2441 | ** | 0.3861 | ** | 0.4682 | ** | 0.3715 | ** | 0.6542 | ** | 0.4973 | * |
| | (3.14) | | (4.56) | | (4.56) | | (3.45) | | (3.34) | | (2.49) | |
| Proportion office buildings | 0.9529 | ** | 0.8728 | ** | 0.8943 | ** | 0.4486 | | -0.0252 | | 0.7134 | |
| | (4.39) | | (2.60) | | (2.96) | | (1.28) | | -(0.04) | | (1.13) | |
| Percent residential | -0.5725 | ** | 0.1832 | ** | -0.5407 | ** | 0.4759 | ** | -0.2884 | * | 0.1195 | |
| | -(13.20) | | (3.38) | | -(9.15) | | (6.63) | | -(2.22) | | (0.84) | |
| Block X block group heterogeneity | | | -1.5309 | ** | | | -1.0166 | ** | | | -0.1095 | |
| | | | -(7.20) | | | | -(3.83) | | | | -(0.20) | |
| ***Spatial lag measures*** | | | | | | | | | | | | |
| Percent black | -1.2075 | ** | -1.4005 | ** | 0.8004 | ** | 0.8823 | ** | -0.7232 | † | -0.9876 | * |
| | -(9.26) | | -(9.74) | | (4.65) | | (4.64) | | -(1.71) | | -(2.23) | |
| Percent Latino | -1.2583 | ** | -1.3281 | ** | 0.3608 | * | 0.4477 | * | 0.3021 | | -0.0032 | |
| | -(9.26) | | -(9.16) | | (2.00) | | (2.28) | | (0.74) | | -(0.01) | |
| Racial/ethnic heterogeneity | 0.2218 | ** | -0.0371 | | 0.4556 | ** | 0.3569 | ** | -0.0462 | | -0.1704 | |
| | (2.61) | | -(0.42) | | (4.06) | | (2.98) | | -(0.20) | | -(0.69) | |
| Concentrated disadvantage | 0.3501 | ** | 0.3458 | ** | 0.1092 | ** | 0.0994 | * | -0.0372 | | -0.0148 | |
| | (11.83) | | (10.55) | | (2.84) | | (2.35) | | -(0.42) | | -(0.16) | |
| Residential stability | 0.2595 | ** | 0.2655 | ** | -0.1428 | ** | -0.1186 | ** | 0.2607 | ** | 0.2329 | * |
| | (9.10) | | (8.22) | | -(3.70) | | -(2.81) | | (2.94) | | (2.48) | |
| Percent vacant units | 5.6908 | ** | 2.9615 | ** | 3.8509 | ** | 0.4790 | | 12.2911 | ** | 11.4894 | ** |
| | (6.77) | | (3.19) | | (3.43) | | (0.39) | | (5.00) | | (4.55) | |
| Population density | 3.2405 | ** | 2.4841 | ** | 3.7333 | ** | 3.7397 | ** | 4.2114 | ** | 3.8262 | ** |
| | (9.14) | | (6.55) | | (7.51) | | (7.51) | | (4.05) | | (3.50) | |
| Intercept | 3.0079 | ** | 0.4544 | | -1.3506 | | -4.7925 | ** | -15.9371 | | 1.9008 | |
| | (3.84) | | (0.52) | | -(1.29) | | -(4.14) | | -(0.02) | | (0.79) | |

*Note: ** p < .01; * p < .05; † p < .10. T-values in parentheses. N = 73,010 blocks across seven cities (Chicago; Cleveland; Columbus; Dallas; Los Angeles; San Francisco; Tucson). Models include fixed effects for cities.*

**Table 3. Multilevel negative binomial regression models for three types of property crime; including block, block group, and spatially lagged measures of demographics, and block and block group measures of land use variables**

| | (1a) | | (1b) | | (2a) | | (2b) | | (3a) | | (3b) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Burglary | | Burglary | | Motor vehicle theft | | Motor vehicle theft | | Larceny | | Larceny | |
| ***Block measures*** | | | | | | | | | | | | |
| Percent black | | | 0.0146 | | | | 0.2593 | ** | | | -0.0047 | |
| | | | (0.23) | | | | (3.16) | | | | -(0.07) | |
| Percent Latino | | | -0.4533 | ** | | | -0.0532 | | | | -0.3728 | ** |
| | | | -(7.29) | | | | -(0.74) | | | | -(5.80) | |
| Racial/ethnic heterogeneity | | | 0.4287 | ** | | | 0.0805 | | | | 0.2587 | ** |
| | | | (5.23) | | | | (0.82) | | | | (2.71) | |
| Concentrated disadvantage | | | -0.2998 | ** | | | -0.2525 | * | | | -0.2831 | ** |
| | | | -(3.46) | | | | -(2.36) | | | | -(2.82) | |
| Residential stability | | | -0.0931 | ** | | | -0.1385 | ** | | | -0.0966 | ** |
| | | | -(6.03) | | | | -(8.20) | | | | -(5.92) | |
| Percent vacant units | | | 1.0632 | ** | | | 0.6118 | ** | | | 0.8028 | ** |
| | | | (13.79) | | | | (6.08) | | | | (9.11) | |
| Population density | | | -2.2328 | ** | | | -1.7439 | ** | | | -1.4877 | ** |
| | | | -(17.26) | | | | -(13.04) | | | | -(8.99) | |
| Percent aged 16 to 29 | | | -0.0609 | | | | 0.1350 | | | | 0.0518 | |
| | | | -(0.69) | | | | (1.37) | | | | (0.53) | |
| Proportion commercial buildings | | | 1.3013 | ** | | | 1.2626 | ** | | | 1.6125 | ** |
| | | | (18.94) | | | | (17.35) | | | | (20.24) | |
| Proportion industrial buildings | | | 0.6578 | ** | | | 0.8456 | ** | | | 0.4584 | ** |
| | | | (8.32) | | | | (10.23) | | | | (5.41) | |
| Proportion office buildings | | | 0.9888 | ** | | | 0.6118 | ** | | | 0.9976 | ** |
| | | | (5.36) | | | | (3.93) | | | | (6.35) | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proportion housing units | | | -0.8364 | ** | | | -1.0268 | ** | | | -1.8396 | ** |
| | | | -(23.31) | | | | -(26.89) | | | | -(42.43) | |
| Case is within 0.1 mile of city border | | | 4.2300 | ** | | | 5.6069 | ** | | | 3.8955 | ** |
| | | | (7.49) | | | | (9.40) | | | | (6.28) | |
| ***Block group measures*** | | | | | | | | | | | | |
| Percent black | -0.0811 | | 0.0492 | | 0.2687 | ** | 0.2227 | * | -0.0596 | | 0.1383 | |
| | -(1.63) | | (0.64) | | (5.14) | | (2.36) | | -(1.10) | | (1.53) | |
| Percent Latino | -0.4602 | ** | -0.0699 | | -0.0175 | | 0.1020 | | -0.4580 | ** | -0.1836 | * |
| | -(9.73) | | -(0.95) | | -(0.37) | | (1.26) | | -(8.97) | | -(2.39) | |
| Racial/ethnic heterogeneity | 0.1073 | * | 0.8596 | ** | 0.1712 | ** | 0.7742 | ** | -0.0273 | | 0.8455 | ** |
| | (2.52) | | (10.23) | | (3.89) | | (8.12) | | -(0.58) | | (9.32) | |
| Concentrated disadvantage | 0.0075 | | 0.0476 | ** | 0.0859 | ** | 0.1053 | ** | -0.0033 | | 0.0258 | |
| | (0.55) | | (2.89) | | (5.94) | | (5.46) | | -(0.22) | | (1.40) | |
| Residential stability | -0.0654 | ** | 0.0028 | | -0.1759 | ** | -0.0612 | ** | -0.1111 | ** | -0.0365 | † |
| | -(4.28) | | (0.14) | | -(11.49) | | -(2.96) | | -(7.02) | | -(1.68) | |
| Percent vacant units | 1.0865 | ** | 0.3414 | * | -0.1501 | | -0.4114 | * | 0.0767 | | 0.0026 | |
| | (7.75) | | (2.10) | | -(1.07) | | -(2.32) | | (0.50) | | (0.01) | |
| Population density | -1.8343 | ** | 0.0289 | | -1.7542 | ** | -0.2620 | * | -1.7005 | ** | 0.0388 | |
| | -(23.88) | | (0.24) | | -(23.36) | | -(2.24) | | -(21.33) | | (0.28) | |
| Percent aged 16 to 29 | 0.3359 | ** | 0.3121 | ** | 0.0505 | | 0.1528 | | 0.2396 | * | 0.4650 | ** |
| | (3.46) | | (2.76) | | (0.50) | | (1.28) | | (2.42) | | (3.73) | |
| Economic inequality | 0.1197 | * | 0.4286 | ** | -0.0349 | | 0.1350 | | 0.0941 | † | 0.5665 | ** |
| | (2.20) | | (5.33) | | -(0.63) | | (1.49) | | (1.67) | | (6.05) | |
| Proportion commercial buildings | 1.1777 | ** | 0.0848 | | 1.7773 | ** | 0.3124 | ** | 2.6079 | ** | 0.1452 | |
| | (12.91) | | (0.85) | | (19.20) | | (2.90) | | (25.14) | | (1.31) | |
| Proportion industrial buildings | 0.8643 | ** | 0.3935 | ** | 1.1980 | ** | 0.3273 | ** | 0.7084 | ** | 0.1485 | † |
| | (12.27) | | (5.08) | | (16.89) | | (4.24) | | (8.97) | | (1.76) | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proportion office buildings | 1.5006 | ** | 0.7699 | ** | 0.6683 | ** | 0.2022 | | 2.0490 | ** | 1.1151 | ** |
| | (7.43) | | (2.95) | | (3.31) | | (0.78) | | (8.95) | | (3.42) | |
| Proportion housing units | -0.0052 | | 0.3395 | ** | -0.2271 | ** | 0.2248 | ** | -1.0070 | ** | 0.0101 | |
| | -(0.13) | | (7.43) | | -(5.62) | | (4.48) | | -(22.86) | | (0.20) | |
| Block X block group heterogeneity | | | -1.5249 | ** | | | -0.8713 | ** | | | -1.4537 | ** |
| | | | -(8.60) | | | | -(4.29) | | | | -(7.15) | |
| ***Spatial lag measures*** | | | | | | | | | | | | |
| Percent black | 0.6928 | ** | 0.6509 | ** | 1.1268 | ** | 0.9964 | ** | 0.4190 | ** | 0.2320 | † |
| | (6.02) | | (5.44) | | (9.35) | | (7.51) | | (3.33) | | (1.89) | |
| Percent Latino | 0.3950 | ** | 0.4451 | ** | 1.0019 | ** | 1.0275 | ** | -0.0764 | | 0.0364 | |
| | (3.29) | | (3.56) | | (8.13) | | (7.71) | | -(0.58) | | (0.28) | |
| Racial/ethnic heterogeneity | 0.7673 | ** | 0.6558 | ** | 0.5435 | ** | 0.4647 | ** | 0.1102 | | 0.0483 | |
| | (9.80) | | (8.48) | | (6.82) | | (5.82) | | (1.28) | | (0.61) | |
| Concentrated disadvantage | 0.0573 | * | 0.0223 | | 0.0348 | | 0.0278 | | 0.0646 | * | 0.0831 | ** |
| | (2.23) | | (0.88) | | (1.32) | | (0.98) | | (2.26) | | (2.94) | |
| Residential stability | -0.0986 | ** | -0.0752 | ** | -0.1916 | ** | -0.1706 | ** | -0.1534 | ** | -0.0967 | ** |
| | -(3.87) | | -(2.73) | | -(7.24) | | -(5.72) | | -(5.46) | | -(3.39) | |
| Percent vacant units | 4.7699 | ** | 3.1569 | ** | -1.6408 | * | -3.3703 | ** | 2.4143 | ** | -0.1897 | |
| | (6.40) | | (4.12) | | -(2.08) | | -(3.86) | | (2.88) | | -(0.24) | |
| Population density | 1.2013 | ** | 0.6716 | * | -1.1248 | ** | -1.7637 | ** | 2.5558 | ** | 1.3298 | ** |
| | (3.57) | | (1.97) | | -(3.33) | | -(5.13) | | (6.90) | | (3.55) | |
| Intercept | 0.4255 | | -1.4651 | * | -4.2292 | ** | -5.9478 | ** | 0.2027 | | -2.5733 | ** |
| | (0.61) | | -(2.03) | | -(5.74) | | -(7.20) | | (0.26) | | -(3.37) | |

*Note: ** $p < .01$; * $p < .05$; † $p < .10$. T-values in parentheses. N = 73,010 blocks across seven cities (Chicago; Cleveland; Columbus; Dallas; Los Angeles; San Francisco; Tucson). Models include fixed effects for cities.*

**Table 4. Summary of results**

| | | Assault | Robbery | Homicide | Burglary | Motor vehicle theft | Larceny |
|---|---|---|---|---|---|---|---|
| Percent black | | | | | | | |
| | Block | + | | | | + | |
| | Block group | + | + | + | | + | |
| | Spatial lag | - | + | - | + | + | |
| Percent Latino | | | | | | | |
| | Block | + | | | - | | - |
| | Block group | + | + | + | | | - |
| | Spatial lag | - | + | | + | + | |
| Racial/ethnic heterogeneity | | | | | | | |
| | Block | + | | | + | | + |
| | Block group | + | + | | + | + | + |
| | Spatial lag | | + | | + | + | |
| Block X block group heterogeneity | | - | - | | - | - | - |
| Concentrated disadvantage | | | | | | | |
| | Block | + | | | - | - | - |
| | Block group | + | + | + | + | + | |
| | Spatial lag | + | + | | + | + | |
| Economic inequality | | | | | | | |
| | Tract | + | + | | + | | + |
| Residential stability | | | | | | | |
| | Block | - | - | | - | - | - |

The scale of ecology

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Block group | + | | + | | - | |
| | Spatial lag | + | - | + | - | - | - |
| Percent vacant units | | | | | | | |
| | Block | + | + | + | + | + | + |
| | Block group | + | | | + | - | |
| | Spatial lag | + | | + | + | - | |
| Population density | | | | | | | |
| | Block | - | - | - | - | - | - |
| | Block group | + | + | + | | - | |
| | Spatial lag | + | + | + | + | - | + |
| Percent aged 16 to 29 | | | | | | | |
| | Block | | | | | | |
| | Block group | | | | + | | + |
| Proportion commercial buildings | | | | | | | |
| | Block | + | + | + | + | + | + |
| | Block group | + | + | | | + | |
| Proportion industrial buildings | | | | | | | |
| | Block | - | | | + | + | + |
| | Block group | + | + | + | + | + | |
| Proportion office buildings | | | | | | | |
| | Block | | + | | + | + | + |
| | Block group | + | | | + | | + |
| Proportion residential | | | | | | | |
| | Block | - | - | - | - | - | - |
| | Block group | + | + | | + | + | |
| Case is within 0.1 mile of city border | | | | | | | |
| | Block | + | | | + | + | + |

*Summary of results from models for six outcomes across seven cities (Chicago; Cleveland; Columbus; Dallas; Los Angeles; San Francisco; Tucson)*

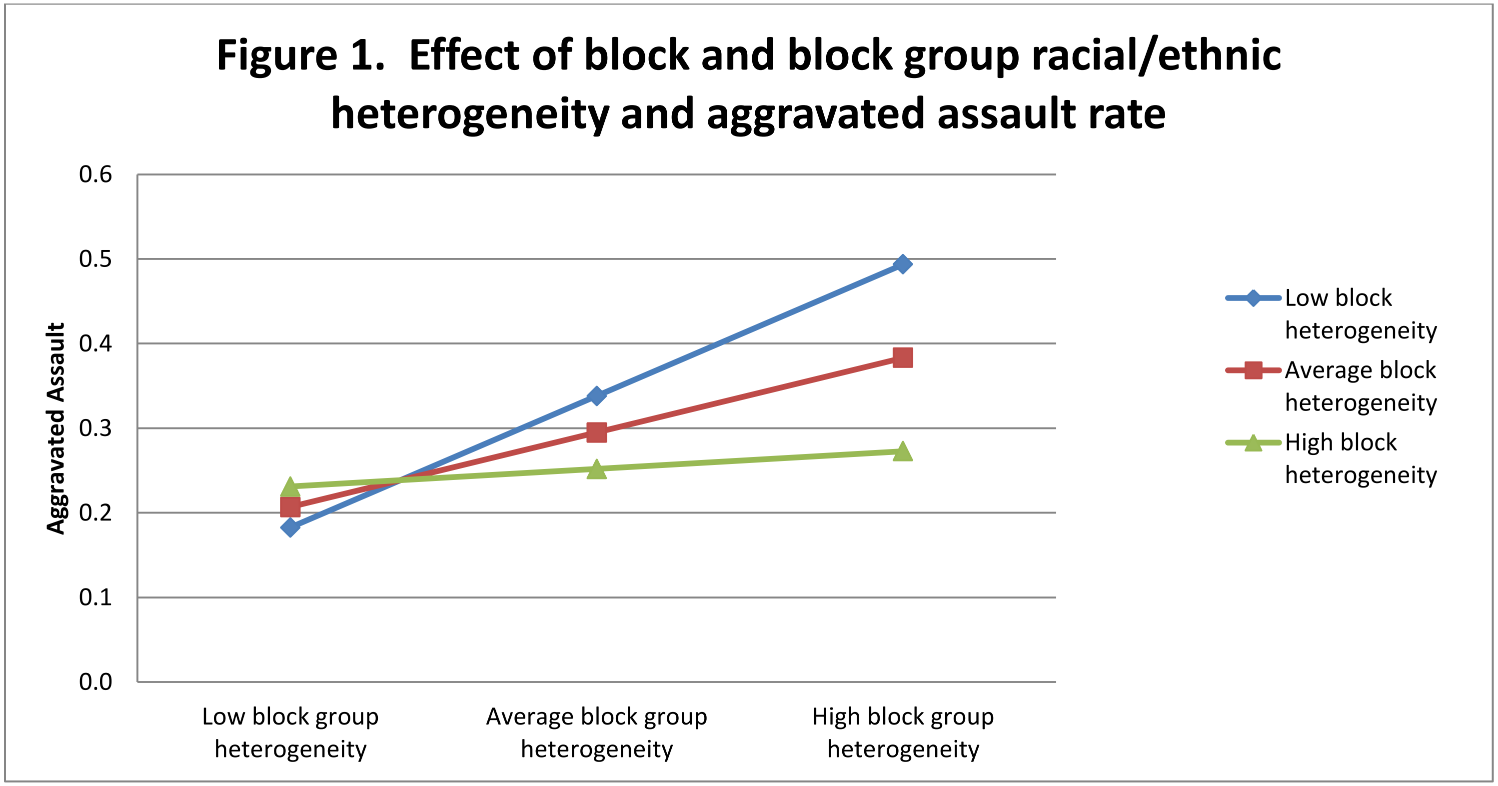
Figure 1. Effect of block and block group racial/ethnic heterogeneity and aggravated assault rate
Aggravated Assault
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Low block group heterogeneity
Average block group heterogeneity
High block group heterogeneity
Low block heterogeneity
Average block heterogeneity
High block heterogeneity

Appendix A: Years of data used in analyses.

| **City** | **Data Type** | **Year** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** |
| Chicago | Crime Data | | X | X | X | | | |
| | Land Use Data | | X | | | | | |
| Cleveland | Crime Data | | | | | X | X | X |
| | Land Use Data | | | | | X | | |
| Columbus | Crime Data | X | X | X | | | | |
| | Land Use Data | | | X | | | | |
| Dallas | Crime Data | X | X | X | | | | |
| | Land Use Data | X | | | | | | |
| Los Angeles | Crime Data | X | X | X | | | | |
| | Land Use Data | | X | | | | | |
| San Francisco | Crime Data | | | | X | X | X | |
| | Land Use Data | | | | | X | | |
| Tucson | Crime Data | | | | | X | X | X |
| | Land Use Data | | | | | | X | |

Note: All other data is from the 2000 Census.

**Online Appendix: Synthetic estimation to impute values for small units of analysis**

When data are available at larger geographic units, but not at smaller geographic units, a technique for imputing values is synthetic estimation for ecological inference (Cohen and Zhang 1988; Steinberg 1979). The synthetic estimation approach relies on the assumption that the relationship between variables at one level of analysis is similar at a different level of analysis, which is certainly not ideal. Nonetheless, whereas researchers oftentimes simply impute values from the larger units to the smaller units assuming homogeneity within the larger units, the synthetic estimation approach is more principled in attempting to build a model to predict such values. A brief treatment of the topic can be found in (Steinberg 1979), whereas a longer discussion of issues involved is contained in (Cohen and Zhang 1988). A more recent treatment of the ecological inference problem can be found in (King 1997).

For ecological inference from larger to smaller units, there are three main issues to confront: 1) the necessity to build a prediction model at the next highest level of aggregation that contains valid values of the variable, and then use this model to predict the values of the variable at the smaller unit; 2) the values of the variable of interest in the smaller units must be constrained to sum to the observed total in the larger unit; 3) the need to account for the uncertainty in this prediction. We adopt such an approach here by building a regression model at the higher level of aggregation, using the coefficient estimates of this model to obtain predicted values in the smaller units, adjusting the imputed values for the smaller units such that they sum to the value in the larger unit they are contained within, and then adding uncertainty to the predicted values based on the uncertainty in the imputation model at the higher unit of analysis.

To demonstrate this approach, we used U.S. Census data and crime data for the city of Los Angeles. We used data aggregated to tracts (N=1,053) to estimate predicted values at the block group level, and compared those to the true values. We used the following four measures in models separately, as well as combined as a measure of *concentrated disadvantage*: 1) percentage single parent households; 2) percentage below the poverty level; 3) average household income; 4) percentage with at least a bachelor's degree. The measure is created using regression scoring of the factor loadings from a confirmatory factor analysis: this measure has a mean of 0 and a standard deviation equal to that of the percent in poverty measure (since this is used to scale the factor). We estimated negative binomial regression models with aggravated assault, and then robbery, as the outcome variables, controlling for standard measures used in the neighborhood context of crime literature.

**Results**

The top half of Table A1 presents the results for several models with aggravated assault as the outcome measure; the bottom half of the table displays similar model results with robbery as the outcome measure. Each row represents the models using a particular variable as the key independent variable of interest, and each column presents the models using a particular imputation strategy (each model also contains all control variables). For example, column 1 displays the various model results when using the actual block group aggregated data. Thus, these are essentially the "gold standard" results as we actually have these various measures aggregated to block groups. Column 2 displays the results when adopting the common strategy of simply imputing the value of a measure for a tract to each of the block groups within that same tract. Column 3 uses our synthetic estimation approach with a single imputation, and column 4 uses our synthetic estimation approach with multiple (5) imputations.

<<<Table A1 about here>>>

We see in row 1 that when using the percent single parent households as the single measure to capture concentrated disadvantage it has a positive, but nonsignificant, effect on aggravated assaults in the true model. In column 2, the approach that simply imputes the mean value of the larger tracts to the block groups results in a much stronger, and significant, effect. Notably, across virtually all of the models we estimated the coefficient estimates are always much *larger* than they are for the true values. Thus, this approach of simply imputing the value from the larger unit into the smaller units within it always results in overestimates of the true relationship in our example dataset. We see in this table that the coefficient for the measure in column 2 is always larger than the corresponding coefficient in column 1 using the "true" measure. These coefficients are typically 50 to 100% larger than the true coefficients, and sometimes much larger than this. The exception is that this approach actually yields a coefficient of the *opposite* sign in the model with single parent households as a covariate in the robbery model. For example, whereas the true measure of percent single parent households did not have a significant effect on aggravated assault in the model using the true measure in column 1, it appears to have a significant positive effect with a coefficient nearly 8 times larger in column 2 when using this mean imputation approach. Clearly, this pattern of results is unsatisfactory.

In column 3 we display the results for our synthetic estimates but only using a single imputation, and column 4 displays the same results when using multiple imputations. Whether using a single or multiple imputations, the coefficient estimates are quite similar, and often reasonably close to the true values (and is typically closer than the approach that simply imputes the mean value to the smaller units). The standard errors are larger for the multiple imputation approach, as expected given that this approach accounts for the uncertainty due to not actually having the measures at the smaller unit of analysis.

## Conclusion

Whereas the synthetic estimation approach is certainly not ideal, we argue that it is preferred to the most common imputation approach of simply imputing the value from a larger unit into the subunits within that unit. We have shown here that this common strategy is quite undesirable. Not only does it produce standard errors that are too small, but in these examples it consistently produced coefficient estimates that were severely upwardly biased. It is also worth emphasizing that another common strategy, simply omitting a variable because it is missing in the smaller units of analysis, is not desirable: this will result in the well-known omitted variable problem for regression analysis, which yields biased estimates. It is therefore essential that researchers directly address this missing data problem. While there are certainly limitations to the synthetic estimation approach, we argue that it is more principled than many of the existing strategies employed by applied researchers.

## References

Cohen, Michael Lee and Xiao Di Zhang. 1988. "The Difficulty of Improving Statistical Synthetic Estimation." Bureau of the Census, Washington, D.C.

King, Gary. 1997. *A Solution to the Ecological Inference Problem*. Princeton, NJ: Princeton University Press.

Steinberg, Joseph. 1979. "Synthetic Estimates for Small Areas: Statistical Workshop Papers and Discussion." Pp. 282 in *National Institute on Drug Abuse Research Monograph Series*, vol. 24. Washington, D.C.: National Institute on Drug Abuse.

## Tables and Figures

Table A1. Using synthetic estimation to create block group level variables based on tract level measures: Negative binomial regression coefficients for aggravated assault and robbery models

| | | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|---|
| | ***Aggravated assault models*** | True value | | Value of larger unit | | Single imputation | | Multiple imputation | |
| (1) | Single parent households | 0.0031 | | 0.0238 | ** | 0.0029 | | 0.0030 | |
| | | (0.0020) | | (0.0034) | | (0.0023) | | (0.0024) | |
| (2) | Poverty | 0.0209 | ** | 0.0325 | ** | 0.0157 | ** | 0.0156 | ** |
| | | (0.0017) | | (0.0018) | | (0.0018) | | (0.0022) | |
| (3) | Average household income | -0.0043 | ** | -0.0077 | ** | -0.0024 | ** | -0.0024 | ** |
| | | (0.0006) | | (0.0007) | | (0.0006) | | (0.0007) | |
| (4) | Education level | -0.0216 | ** | -0.0293 | ** | -0.0096 | ** | -0.0100 | ** |
| | | (0.0019) | | (0.0019) | | (0.0018) | | (0.0021) | |
| (5) | Concentrated disadvantage index | 0.0445 | ** | 0.0506 | ** | 0.0222 | ** | 0.0223 | ** |
| | | (0.0030) | | (0.0027) | | (0.0028) | | (0.0030) | |
| | | | | | | | | | |
| | ***Robbery models*** | | | | | | | | |
| (1) | Single parent households | -0.0093 | ** | 0.0163 | ** | -0.0066 | * | -0.0058 | * |
| | | (0.0024) | | (0.0042) | | (0.0029) | | (0.0029) | |
| (2) | Poverty | 0.0201 | ** | 0.0374 | ** | 0.0196 | ** | 0.0196 | ** |
| | | (0.0021) | | (0.0022) | | (0.0022) | | (0.0024) | |
| (3) | Average household income | -0.0021 | ** | -0.0090 | ** | -0.0029 | ** | -0.0026 | ** |
| | | (0.0007) | | (0.0009) | | (0.0008) | | (0.0009) | |
| (4) | Education level | -0.0153 | ** | -0.0222 | ** | -0.0063 | ** | -0.0065 | * |
| | | (0.0023) | | (0.0024) | | (0.0022) | | (0.0027) | |
| (5) | Concentrated disadvantage index | 0.0304 | ** | 0.0515 | ** | 0.0184 | ** | 0.0191 | ** |
| | | (0.0038) | | (0.0035) | | (0.0035) | | (0.0036) | |

*Note: ** p < .01; * p < .05; † p < .1. T-values in parentheses. All models control for: percent vacant units, percent owners, percent African American percent Latino, racial/ethnic heterogeneity, population density, and the percent aged 16 to 29. N=2,759 block groups*