The Relationship between Voting and Crime: A Neighborhood-Level Analysis

Starr Moss

University of Wisconsin-Madison Department of Geography

550 North Park Street, Madison WI 53703

e-mail: starrmoss1@gmail.com

Abstract

This article examines the relationship between voting behavior and crime. Studies by Wolfinger (1980), Knack (1992), Coleman (2002) and Gerber (2004) have demonstrated that voting is a socially-conforming behavior. Since voting may indicate participation in broader norm-governed behavior, researchers have examined whether or not there is a relationship between voter turnout and crime. Coleman (2002) showed that voter turnout and crime rate exhibit a parabolic relationship at the county and state level, but to date, no studies have explored the relationship between voter turnout and crime at the neighborhood level. In addition, no study has examined the spatial structure of the relationship between voter turnout and crime. This research attempts to address these gaps in the existing research.

1. Introduction

Why do people vote? Casting a ballot on election day is not just a way to show support for a particular candidate, it also embeds an individual into collective society. Knack (1992) argues that voting is a social norm, or a *collectively understood rule that prescribes human behavior*. Coleman (2002) theorizes that voting is an example of social conformity, which he defines as the *alignment of people's thinking or behavior with a societal or group norm*. Coleman (2002) also discovered a parabolic relationship between voter turnout and crime rate at the state and county level. However, few studies have examined whether this relationship is true at the neighborhood level.

In order to see whether the patterns observed by Coleman (2002) might be reproduced at a much finer spatial scale, the interaction between voting and crime needs to be studied at the neighborhood level. Thus, my research focuses on how voter behavior influences crime rates at a neighborhood level. This research aims to show the relationship between voting behavior and crime at the neighborhood level, and the spatial structure of this relationship.

2. Social Conformity and Voting

In order to understand voting as a socially conforming behavior, it is important to understand what motivates people to vote. Appealing to common sense, Riker (1968) points out that the act of voting is fundamentally irrational. A rational individual has little incentive to vote, knowing that the costs of voting – registering, researching candidates, travelling to the polls, and waiting in line – exceeds the benefits – the small chance that casting their ballot makes a difference. Yet, as Blais (2000) observes, the majority of people vote in major elections, even when they recognize that their ballot's influence is inconsequential. This begs the question: why *do* people vote?

The answer – suggested by Wolfinger et al (1980), Knack (1992), Gerber (2004), and Coleman (2002) – is that voting is "norm-governed behavior", and thus the decision to vote is strongly influenced by external social pressure. Here, for the sake of clarity, I will return to Cialdini's (1998) concept of a social norm as a collectively understood rule that prescribes human behavior. As Riker (1968) and Blais (2000) have pointed out, voting is irrational at the individual level – a person has little intrinsic motivation to vote knowing that their vote will not likely decide the election. Knack (1992) explains this phenomenon by arguing that the social obligation and "duty" to vote is what provides the incentive to vote, rather than the prospect of determining an election's outcome. In other words, voting appears irrational as an individual choice, but when seen as a social norm, an individual's motivation to vote makes much more sense.

Knack (1992) characterizes voting as a social norm. He theorizes that voting is not only a civic institution but a *societal* institution, and people are more pulled by the act of voting itself rather than actually influencing the election. Knack (1992) cites a study by the National Election Board that show a majority of people in a Wisconsin poll believe it's important to vote even if they know their party cannot win. Dennis (1970) explains this phenomenon, suggesting that "the average member of the public will more likely have internalized the norms of electoral participation than those of partisan competition...Voting and elections are 'us'; parties are 'them.'" Knack (1992) also argues that while voting may be motivated by a civic duty or ethical obligation, the value placed on voting is still transferred and reinforced socially. He bolsters this case by presenting evidence that social sanctions provide strong voting incentives. As an example, he cites Alderman's (1983) study that shows 41% of regular voters cited pressure from family and friends as a reason they voted. Knack (1992) also points to evidence that spousal enforcement of voting norms is a major contributor to voting turnout in marriages, which may account for the fact that married individuals vote more than non-married.

While Knack (1992) and Wolfinger (1980) present convincing evidence that voting is a salient social norm, Gerber's (2004) work goes one step further by directly linking voting conformity to perceived social norms about voting. Asch (1956), Crutchfield (1955) and Deutsch et al (1955) argue that conformity is motivated by either the desire to make a factually correct decision, or is motivated by *normative influence* – the desire to go along with the group and its prescriptive norms. Gerber (2004) supports the latter explanation, showing in his study that individuals' tendency to vote can be highly motivated by a desire to conform with the majority and avoid derision. He finds that people are 7% more likely to vote when they perceive that voter turnout was high, compared to low. Furthermore, Gerber (2008) finds that voter turnout is directly influenced by social pressure. Applying different levels of social pressure by threatening to expose an individual to their neighbors if they do not vote, Gerber (2008) finds that voter turnout increases proportionally to increased levels of social pressure.

3. Voting and Crime

Having established that voting is a viable measure of conformity to a social norm, I will turn next to whether voting, as a potentially norm-governed behavior, has an effect on crime. If voting demonstrates a willingness to adopt the voting norm, does it show willingness to adhere to other norms, such as not breaking the law? In order to justify this question, I address whether conformity to voting might be indicative of broader conformity to social and civic-minded norms that may influence crime rate. While there is no clear answer to this question, there is evidence (albeit limited) that conformity in voting may lead to conformity to other social norms. Knack (1992) links individual voting behavior to larger-scale "socially cooperative behavior" in terms of social and civic engagement. He finds that those who vote are more likely to be involved in various neighborhood institutions like Parent Teacher Associations (PTA), charitable organizations, and other political groups. He also finds that voters, compared to non-voters, are more likely to respond to the U.S. census. Finally, Knack cites a District of Columbia study by Tyler (1990) that shows the correlation coefficient between turnout and crime is -.43 at the state level, and -.30 at the neighborhood level. With this evidence, Knack (1992) suggests that voting may fall under a broad category of "norm-governed" behavior.

Uggen et al. (2004) echoes the idea that voting may be linked to broader social norms, offering evidence that voting may be a mechanism for civic and social integration. Uggen et al (2004) hypothesizes that voting creates "reciprocal obligations" between an individual and society, and that these bonds deter antisocial criminal behavior. He notes that voting is negatively correlated with criminal recidivism – an individual's repeating of criminal behavior. Using 757 respondents who self-reported past arrests, Uggen et al (2004) found that compared to people who voted in the 1996 presidential election, those who abstained had higher rates of criminal recidivism. However, Uggen et al. (2004) only uses voting in the 1996 election as a basis for the study, and employs a limited sample population. While he presents a convincing case that voting is indeed linked to lower rates of criminal recidivism in individuals, it remains unclear whether this pattern would hold true at the macro-level. A geographic analysis, employing a larger sample population over a longer period of time, is needed to provide insight into how patterns between electoral participation and criminal behavior changes over time and space.

Coleman (2002) gets closer to this question. He pinpoints a clear relationship between voter turnout and crime rates over a period of time, though his research is generalized to the county and state level and therefore may miss nuances in how the relationship between voting and crime changes at the community level. Coleman (2002) uses voter turnout as a measure of social conformity in order to assess conformity's effect on crime rates. In a study of state-wide voter turnout for presidential elections, Coleman (2002) finds that the relationship between voter turnout and crime is parabolic; high voting conformity is negatively correlated to crime, as is low voting conformity. Crime rate peaks at voter turnout "entropy", where voter turnout is roughly 50%. Coleman (2002) argues at roughly 50% voter turnout, social conformity to the voting norm is absent.

The existing studies referenced above show that voting may likely be a measure of social conformity, and suggest that at least in some instances, there is an inverse relationship between voting and crime. The relationship between voting and crime has been examined at the individual level by Uggen (2004), and at the county and state level by Coleman (2002), yet few studies have examined how voting and crime interact at the neighborhood level.

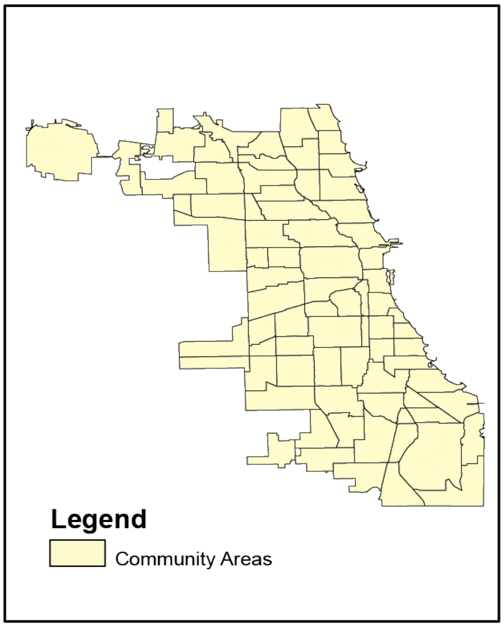
The need for neighborhood-level research on the relationship between voting and crime is twofold. First, the county and state – the units of analysis used by Coleman (2002) – might not be as effective as transmitting social norms and enforcing social conformity compared to the neighborhood. It is likely that some people identify more as citizens of the particular neighborhood they live in rather than the state or county they reside in. Because the neighborhood might better encapsulate notions of social norms and conformity, it may be a more appropriate unit of analysis. Second, it is possible that the patterns observed by Uggen (2004) and Coleman (2002) are subject to the *scale effect* of Openshaw's (1983) modifiable unit area problem (MAUP) – which states that observable patterns vary based on the geographic scale of analysis. In other words, results obtained at one geographic scale might not hold true for a different geographic scale. In order to determine whether or not the results observed by Knack (1992), Uggen et al (2004) and Coleman (2002) would be replicated at the neighborhood level, my research will focus how voter behavior and crime rate interact at the neighborhood scale.

4. Study Area and Key Variables

4.1 Study Area

Chicago, Illinois was selected as the study area. For a spatiotemporal assessment of the relationship between voting and crime at the neighborhood level, Chicago is ideal. Being one of the largest cities in the United States – both geographically and population-wise – Chicago has a large enough spatial extent to show meaningful variation in voter turnout and crime rate. Chicago also keeps a rich database of publicly available crime data dating back to 2001. These data are often hard to obtain. The availability of these data sets for Chicago makes it possible to conduct this research.

Chicago's 77 community-areas (neighborhoods) served as enumeration units for aggregated voting and crime data and thus the community-area will be the basis for the analysis (Map 1).



Map 1: Chicago community-areas

The benefits of a community-area level analysis are two-fold. Though performing the analysis at the precinct level seems intuitive because voter turnout is reported by precinct, a precinct level analysis becomes problematic because block-group boundaries – the spatial units for which the American Community Survey (ACS) reports socioeconomic data – do not fall wholly within precinct boundaries. However, blocks and block-groups fall wholly within community-areas since community-areas are drawn at the block level. Therefore, a major benefit to a community-area level analysis is that block-group level ACS data that is necessary for the analysis is easily aggregated up to the community-area level. Second, community-areas are relatively large geographic units of analysis and have substantially greater populations than smaller units such as voting precincts, census tracts, or census block groups. Using units with greater populations may help avoid large variations in the dependent variables, a trend that Coleman (2002) observed when comparing the crime rates for sparsely populated counties and more heavily populated counties. If the unit of analysis has too small a population, the analysis may be overly sensitive to variations in the independent variables.

4.2 Key Data and Variables

Voter turnout was the metric for assessing voting behavior. Voter turnout is defined as total votes cast divided by voting age population (VAP) in a given community-area. Using VAP helped avoid skewed voter turnout results, as certain community-areas have higher populations of young people who cannot vote. It is common, particularly in the political redistricting process, to define voter turnout as total votes cast divided by citizen voting age population (CVAP) in order to account for those who are not legally able to vote due to non-citizen status. However, CVAP data is based on estimates over 5-years, and thus may not be appropriate when making estimates for single years 2010 and 2012, respectively. Therefore, this analysis does not take CVAP data into consideration.

Turnout was gathered for years 2010 and 2014 in the form of aggregated totals of ballots cast per race, per precinct. Since voter turnout is reported at the precinct level, these voter turnout totals were aggregated from roughly 2,000 precincts up to the community-area, ultimately giving each community-area a value for turnout. For each of the above years, the election for Cook County Clerk – a race in which all Chicago residents were able to vote in – was used in order to assess turnout. Each precinct in Chicago had a voter turnout value, thus the entire study area was represented.

To assess crime rates, I primarily used *violent crimes,* which are crimes that involve "*force or the threat of force*" per the Federal Bureau of Investigation's (FBI) Uniform Crime Reporting (UCR) program. Specifically, I focused on homicide, rape, robbery, and assault, which the UCR deems the most serious violent crimes. I also added burglary to the analysis, since Coleman (2002) used Burglary in his model. Using the same five crime types that Coleman (2002) used made it possible to compare my results against his. This process is detailed further in the next section.

This research operates under the assumption that social conformity influences violent crime to some degree. Though Glaeser et al (1996) have found evidence that social conformity – at least in a state-level analysis – has less of an influence on murder and rape than other types of crimes, Coleman (2002) points to research that suggests social conformity does have an effect on certain violent crimes. With this in mind, it would be expected that if voter turnout is indeed a barometer of social conformity, a community-area's crime rate would vary depending on its level of voter turnout.

Fortunately, geo-referenced Chicago crime data dating back to 2000 is publicly available from the City of Chicago (Figure 2). Police reports for each crime type (homicide, rape, robbery, assault, and burglary) with latitude/longitude coordinates referencing where the crime occurred allows for easy importation of crime data into a GIS, where it is converted to point-data and aggregated up to the community-area level. For each crime type, the crime rate is defined as the number of incidents of a particular crime type in a given community-area, divided by the total population of a given community-area. This number is then multiplied by 10,000 to obtain a per-capita rate. Chicago crime data for years 2010 and 2014 was used to correspond to voter turnout for years 2010 and 2014.

Finally, there was a need to control for other common predictors or "explanatory variables" of crime. Land, McCall, and Cohen (1990) present other variables that are used as common crime predictors, including: socioeconomic status, racial composition, population density, median family income, % families under the poverty line, and % families unemployed, all of which are obtained from the U.S. census and American Community Survey (ACS). The variables listed above form the basis for the explanatory variables used in this research, but the particular explanatory variables chosen varied depending on the analysis used. This is further detailed in the following sections. The complete list of candidate explanatory variables used in this research is (with variable name indicated in parenthesis): percentage black (%_Black), percentage white (%_White), percentage Asian (%_Asian), percentage Latino (%_Latino), households with no male present (%_NoMaleHH), single-parent households (%_SingleHH), percentage of population aged 15-34 (%_Age15to34), percentage of divorced males (%_DivorcedMales), percentage of population under 25 with a high-school degree or less (%_HSorLess), percentage of labor force unemployed (%_Unemployment), percentage of households in poverty (%_HHPoverty), percentage of families in poverty (%_FamilyPoverty), percentage of population renting (%_Renter), % of population who pay $600 or less in monthly rent, (%_Rent600Less,) log of total population (Log_Pop), log of population density (Log_PopDensity), and log of income (Log_Inc).

5. Statistical Analysis

5.1 Overview

In order to evaluate the relationship between voting and crime at the community-area level and the spatial pattern of this relationship, the analysis and methodological approach is informed by two questions. First, can Coleman's (2002) results can be replicated at the community-area level? In other words, will the same results emerge even though the scale of analysis changes? The second question is whether another model might better explain the data in this research. This begs two follow-up questions: in such a model, what is the relationship between voter turnout and crime, and what is the spatial pattern of this relationship?

5.2 Replicating Coleman's (2002) study

In order to answer question #1 – whether Coleman's (2002) results at the state and county level could be reproduced at the community-area level – I used the same independent and dependent variables utilized in his study. Following Coleman's (2002) model, I took the log of the per-capita rate of each crime type. Log homicide, log rape, log assault, log robbery, and log burglary were used as dependent variables. The independent variables used were also repeated from Coleman's (2002) model: %_Turnout, turnout^2, % single parent households, log income, log population density, % males divorced, % population aged 15-19, and % population in poverty. In order to determine the impact of the %_Turnout variable, two Ordinary Least Squares (OLS) regression models were run for each crime type. The first OLS model regressed a particular crime type on all independent variables listed above, but omitted the %_Turnout variable. The second OLS model was exactly the same as the first OLS model, but added the %_Turnout variable as an independent variable.

5.3 Developing an Optimal Model

This section deals with question #2 in section 5.1. Though repeating Coleman's (2002) study at the community-area scale might not reveal a strong relationship between voter turnout and crime rate, it is possible that a different scale of analysis requires a different set of explanatory variables. To address this issue, stepwise regression was used to regress each crime type on all possible independent variable combinations and select an independent variable subset that produced the highest quality model according to the Akaike Information Criterion (AIC). Each crime type was regressed on the same 17 candidate variables: %_Black, %_White, %_Asian, %_Latino, %_NoMaleHH, %_SingleHH, %_Age15to34, %_DivorcedMales, %_HSorLess, %_Unemployment, %_HHPoverty, %_FamilyPoverty, %_Renter, %_Rent600Less, Log_Pop, Log_PopDensity, and Log_Inc. The optimal model for each crime type varied (Figure 1).

Log Homicide	Log Rape	Log Assault	Log Robbery	Log Burglary
%_White	%_White	%_White	%_Black	%_White
%_Black	%_Black	%_Latino	%_Asian	%_Asian
%_Asian	%_Asian	%_SingleHH	%_Latino	%_MalesDivorced
%_Latino	%_NoMaleHH	Log _Pop	%_Unemployment	%_HSorLess
%_HSorLess	%_Rent600Less	Log_PopDens	Log_Pop	%_HHPoverty
%_HHPoverty	Log_Pop	%_Age15_34	Log_Inc	%_FamPoverty
%_Renter	Log_Inc			%_Renter
Log_PopDens	%_Age15_34			Log_Pop
%_Age15_34				Log_PopDens
				%_Age15_34

Figure 1: Optimal variable subsets determined by stepwise regression, for year 2010

Once an optimal model was identified for each crime type, OLS regression was performed twice for each crime type. Using the same approach detailed in the previous section, the first OLS model used the optimal variable subsets identified by stepwise regression as independent variables for each crime type (Figure 1). The second OLS model was exactly the same as the first OLS model, but added the %_Turnout variable as an independent variable. This made it possible to assess the impact of the %_Turnout variable on the optimal model for each crime type.

5.4 Geographically Weighted Regression (GWR)

Though OLS regression can provide insights about the relationship between voter turnout and crime, there are limitations to the OLS approach. Linear regression models such as OLS produce one estimate for each variable, assuming that the relationship between dependent variables and independent variables is constant over space. A relationship that does not change over space is referred to by Fotheringham (2002) as "*stationary*". In this sense, linear regression techniques such as OLS are referred to as "global models", with an estimate being the same at every location.

Geographically Weighted Regression (GWR) is an extension of the basic linear regression model with one fundamental difference: it assumes that the relationship being modeled varies over space. According to Fotheringham (2002), GWR is best suited to model relationships that exhibit *non-stationarity*, which means that they change over space. Fotheringham (2002) argues that while many physical processes are fixed over space, social processes often vary, and GWR can be a valuable tool to explore the spatial variation of these processes. Voting and crime, both of which are social processes, thus are ideal candidates for GWR analysis. The basic equation for GWR is:

$$y_i = \beta_0(u_i, v_i) + \Sigma_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i$$

Where (u_i, v_i) are the geographic coordinates of point i and $\beta_k(u_i, v_i)$ represents "the realization of the continuous function $\beta_k(u_i, v_i)$ at any point i" according to Fotheringham (2002). In essence, GWR model can estimate parameter values at all points in space. This is done by a weighting function that weights points closer to the point of estimation more heavily than points further away.

5.5 GWR Analysis of Voter Turnout and Crime

As mentioned previously, the OLS results can offer limited insight into the relationship between voter turnout and crime rate at the community-area level. However, in order to understand the significance of voter turnout across space and the spatial structure and variation of voter turnout's relationship to crime rate, GWR will be used.

Fotheringham (2002) identifies two main issues associated with GWR: the selection of a weighting scheme and bandwidth. Weighting schemes can be either fixed or adaptive, though the basic idea behind either weighting scheme is to use data points surrounding the point of regression in order to estimate a value at the point of regression. For this analysis, the centroid of each community-area is used as the point of regression. In a fixed scheme, a region with a particular distance or *bandwidth* is prescribed around the point of regression. Data points falling within this specified distance will be used to estimate the value at the point of regression. The weight of a data point is larger when it is closer to the point of regression, and smaller when it is further away. In an adaptive scheme, the bandwidth prescribed around the point of regression changes based on fluctuations in the data. In areas with fewer data, the bandwidth will be greater to include more data, and in areas with more data, the bandwidth will be smaller.

GWR 4.0, a software package for geographically-weighted modelling, was used to perform a GWR analysis of voter turnout at the community-area level. An adaptive Gaussian weighting scheme was employed. Instead of allowing the bandwidth to vary (a measure of linear distance), the adaptive Gaussian scheme allows the amount of "nearest neighbors" – data points used in making an estimation at the point of regression – to vary. In this case, the adaptive Gaussian weighting scheme allows the number of community-areas taken into consideration when making an estimation to vary.

6. Results

6.1 Reproduction of Coleman's Work

By repeating Coleman's (2002) analysis at the community-area level, two clear differences emerge. First, Coleman (2002) found that %_Turnout was statistically significant variable for all crime types. At the community-area level, for the year 2010, %_Turnout was only statistically significant for assault and robbery. Second, while Coleman's (2002) study found that the crime rate and %_Turnout exhibited a parabolic relationship, with crime rate peaking at roughly 50% turnout, the data in this research demonstrated a linear relationship rather than quadratic (Figures 2 – 6). The R^2 value suffered considerably when the data was fit to a quadratic model that regressed crime rate on %_Turnout and turnout squared.

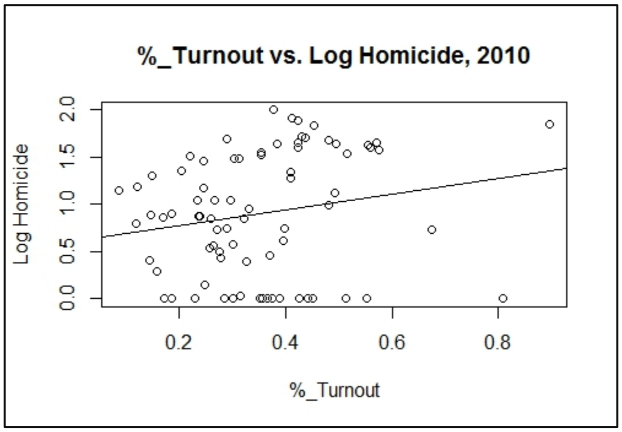


Figure 2: %_Turnout vs. Log Assault

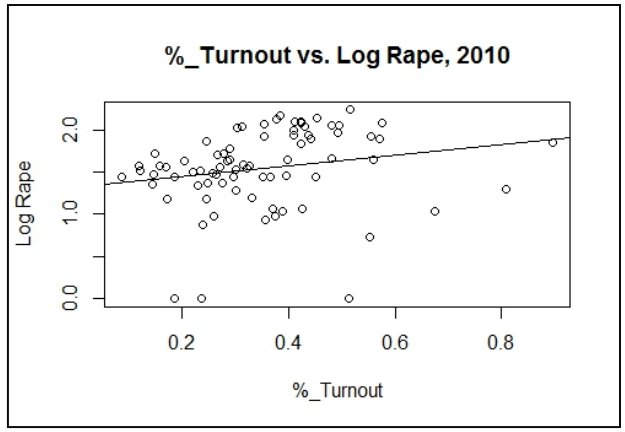


Figure 3: %_Turnout vs. Log Robbery

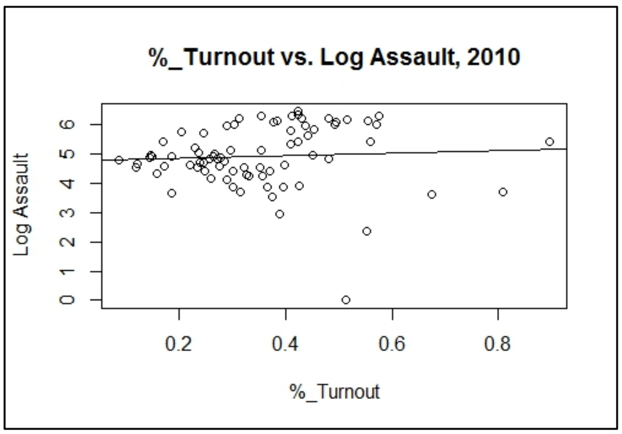


Figure 4: %_Turnout vs. Log Assault

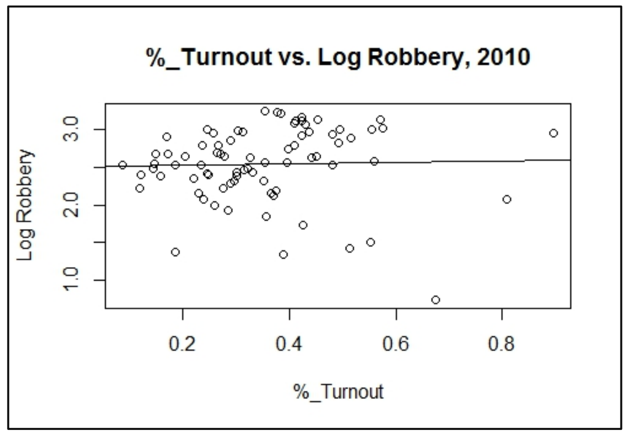


Figure 5: %_Turnout vs. Log Robbery

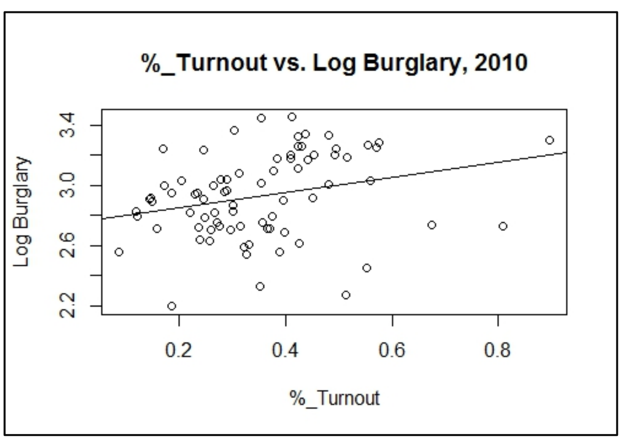


Figure 6: %_Turnout vs. Log Burglary

For each crime type, the R^2 value increases slightly when %_Turnout is added to the model, with homicide seeing the largest increase (Table 1). However, the introduction of the %_Turnout variable clearly does not make a significant impact on the explanatory power of the model.

	Log Homicide	Log Homicide +Turnout.	Log Rape	Log Rape +Turnout.	Log Assault	Log Assault +Turnout.	Log Robbery	Log Robbery +Turnout.	Log Burglary	Log Burglary +Turnout.
R^2	.698	.710	.666	.671	.734	.748	.719	.740	.725	.727

Table 1: R^2 for Log Crime Types, using Coleman's (2002) variables, 2010

6.2 Results from Optimal Model

While each crime type's optimal model yielded a higher R^2 compared to Coleman's model, indicating an improvement in the explanatory power of the optimal models identified by stepwise regression, the effect of the %_Turnout variable was still negligible across all crime types (Table 2). Interestingly, homicide saw the largest increase in R^2 with the addition of the %_Turnout variable, while rape showed zero increase in R^2 with the addition of the %_Turnout variable.

	Log Homicide	Log Homicide +Turnout	Log Rape	Log Rape +Turnout	Log Assault	Log Assault +Turnout	Log Robbery	Log Robbery +Turnout	Log Burglary	Log Burglary +Turnout
R^2	.767	.772	.720	.720	.910	.917	.909	.912	.633	.638

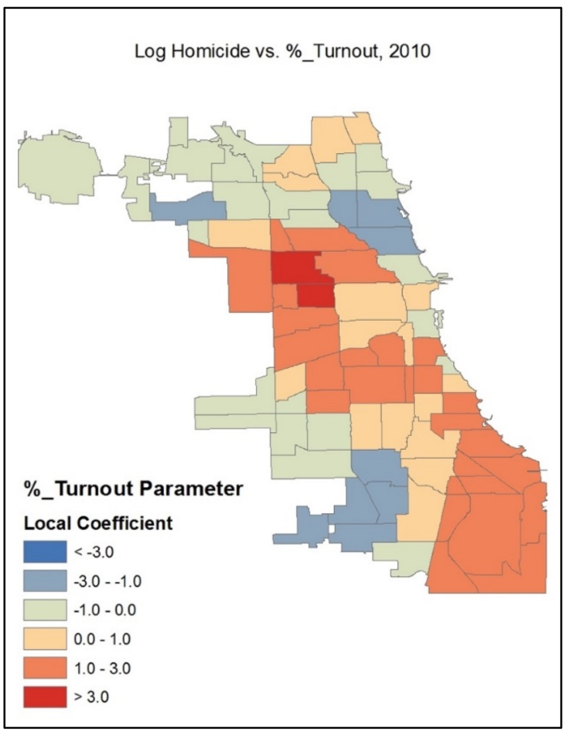
Table 2: R^2 for Log Crime Types, using optimal model from stepwise regression

6.3 Observed Spatial Patterns

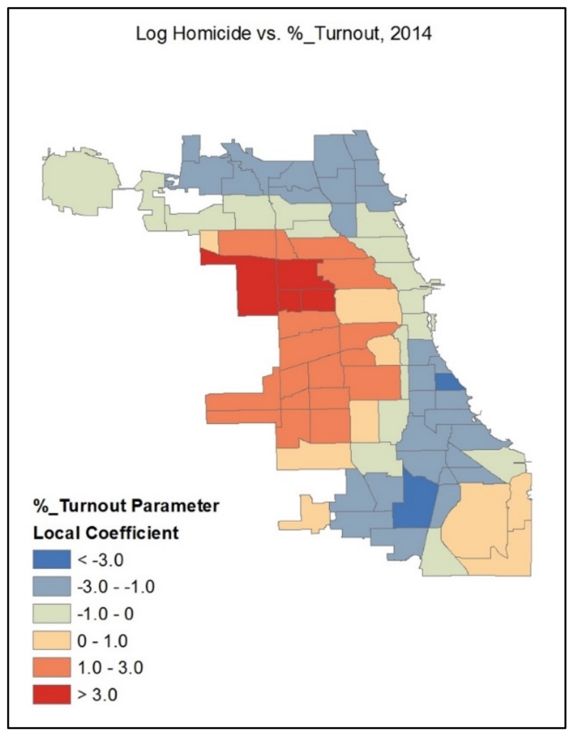
One of the most obvious questions raised with GWR analysis is whether the observed patterns between voter turnout and crime are random. In order to answer this question, the GWR analysis was performed for the year 2010 and 2014 in order verify whether the pattern observed in 2010 was also observed in 2014 (Maps 2 - 9).

6.3.1 Homicide

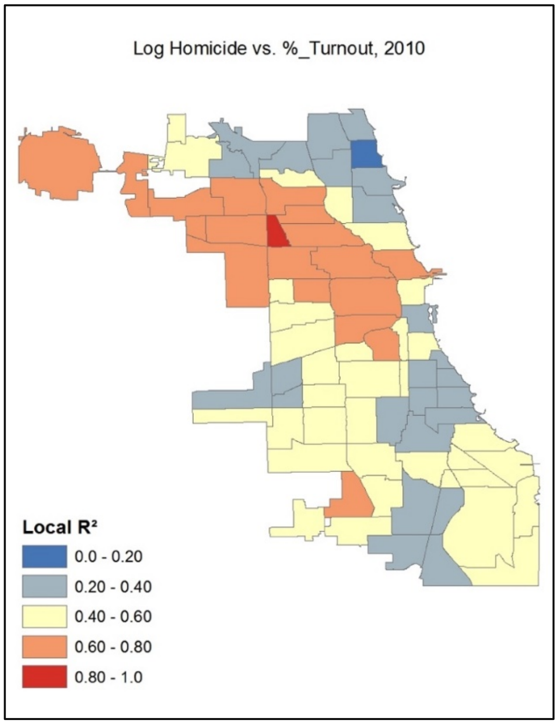
For homicide, an interesting spatial pattern emerges. In 2010 and 2014, there is a high positive coefficient for the %_Turnout variable in an area on the west side of Chicago. For both years, this area on the west side also has a higher local R^2 value compared to the rest of Chicago. A band of community-areas with negative coefficients for %_Turnout can be observed along the north side and south side of Chicago – a pattern that is particularly pronounced in 2014. Examining the %_Turnout coefficients and local R^2 indicates that the association between voter turnout and homicide rate varies from negative or positive depending on spatial location. Finally, the residuals of the homicide model are randomly distributed, indicating a reasonably well-specified model.



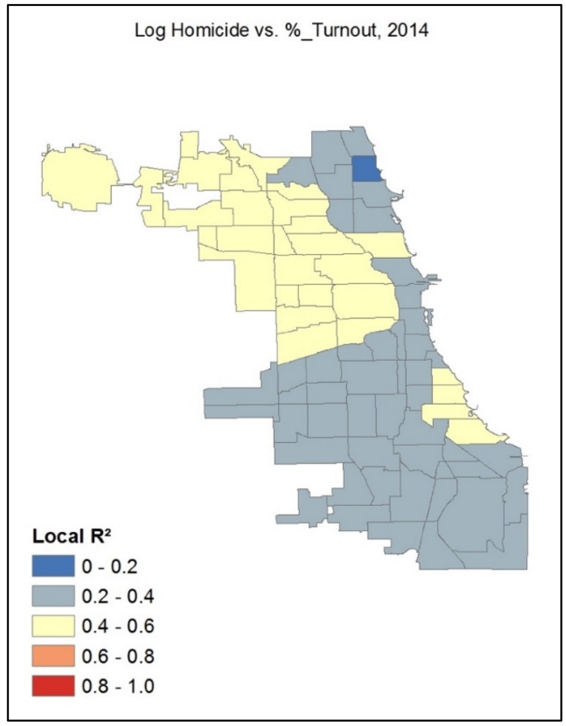
Map 2: Local Coefficients, Log Homicide vs. %_Turnout, 2010



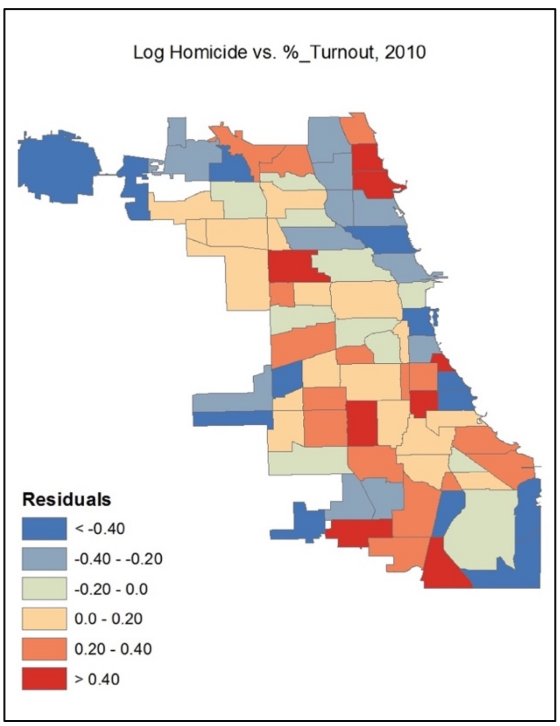
Map 3: Local Coefficients, Log Homicide vs. %_Turnout, 2014



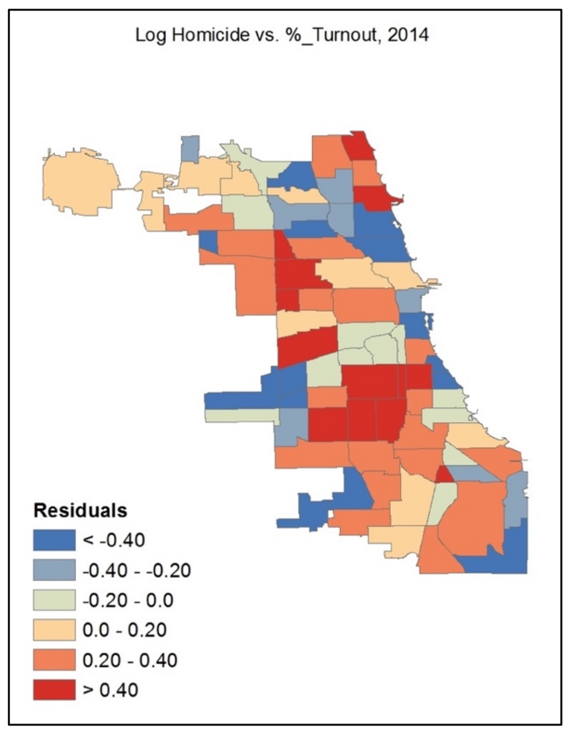
Map 4: Local R^2, Log Homicide vs. %_Turnout, 2010



Map 5: Local R^2, Log Homicide vs. %_Turnout, 2014



Map 6: Residuals, Log Homicide vs. %_Turnout, 2010



Map 7: Residuals, Log Homicide vs. %_Turnout, 2014

6.3.2 Rape

As with homicide, rape also exhibits a unique spatial pattern, with high coefficients for %_Turnout clustering broadly in central Chicago and low coefficients for %_Turnout clustering along the city's north side and south side. However, compared to the homicide model, rape and %_Turnout did not have as strong of a negative association on Chicago's south side. The local R^2 value, as with the homicide model, is consistently highest in an area on the northwest side of Chicago and lowest on the southwest side, indicating spatial variation in the explanatory power of the rape model. As with homicide, these patterns remain consistent for years 2010 and 2014. The residuals are randomly distributed.

6.3.3 Assault

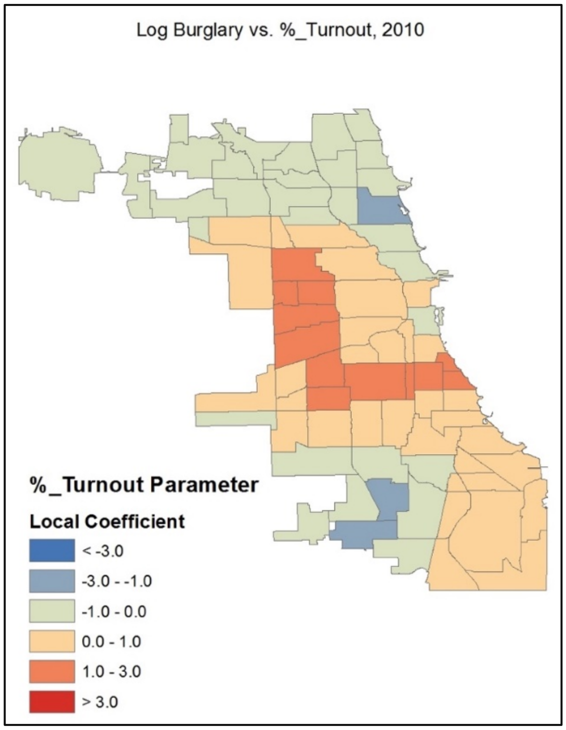
Generally speaking, in terms of the %_Turnout coefficient, assault follows the same trend seen in the homicide and rape models. The %_Turnout coefficient is high in central Chicago and low along the north and southwest portions of the city. For both 2010 and 2014, the highest R^2 (80% or greater) is observed in a cluster of community areas on the northwest side. Interestingly, from 2010 to 2014, much of the southwest side sees a significant rise in local R^2 values. Again, the residuals are randomly distributed.

6.3.4 Robbery

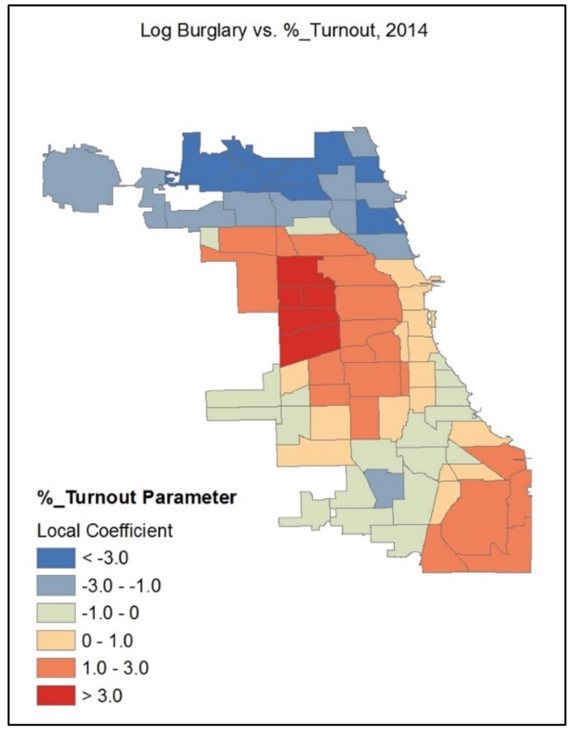
With respect to robbery, the general pattern holds. Again, the %_Turnout coefficient is highest in a broad area of central and west Chicago. From 2010 to 2014, the %_Turnout coefficient increases in a few select community-areas on Chicago's west side. As with homicide, rape, and assault, the %_Turnout coefficient is lowest along the north and southwest sides of the city. The residuals are randomly distributed.

6.3.5 Burglary

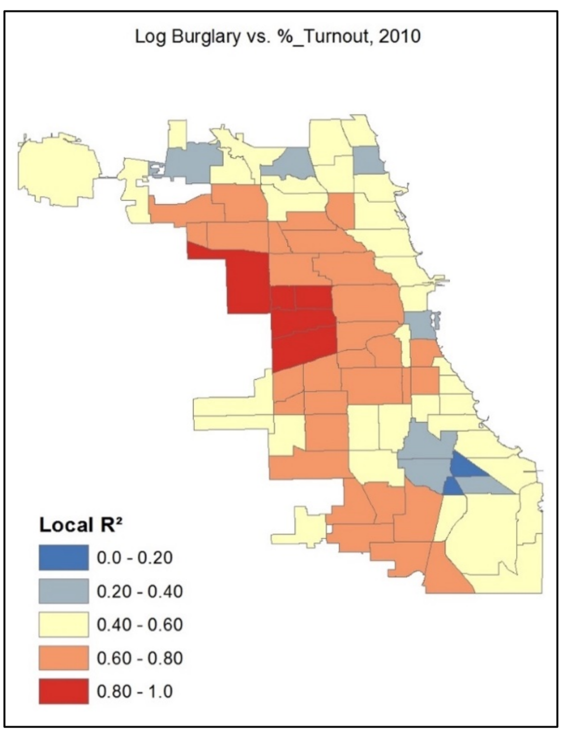
With respect to burglary, the pattern is somewhat unique. While coefficients for %_Turnout are still generally highest on the west side of Chicago, in 2010 there is a distinct "L" shape of community-areas with %_Turnout coefficient values between 1.0 and 3.0. It is not clear why this is. In addition, the %_Turnout negative coefficient values on the north side of Chicago decrease from 2010 to 2014, indicating that the negative association between turnout and burglary here becomes even stronger over this time period. As in prior models, the residuals are randomly distributed spatially.



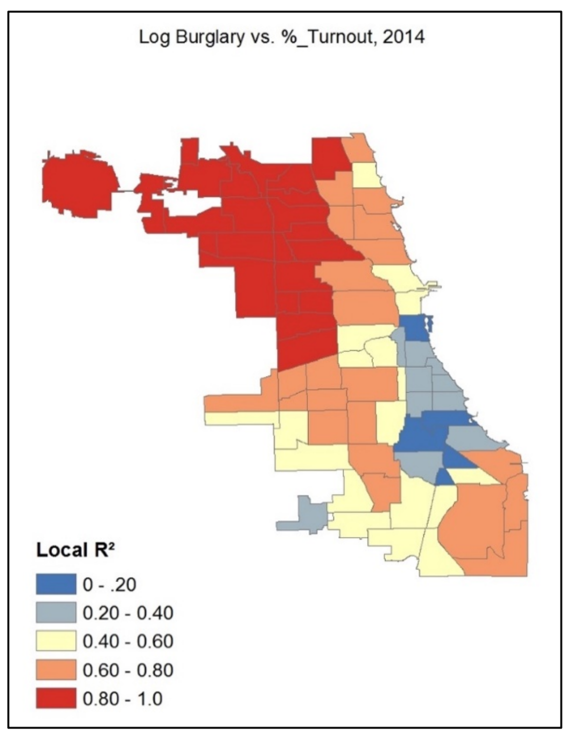
Map 8: Local Coefficients, Log Rape vs. %_Turnout, 2014



Map 9: Local Coefficients, Log Rape vs. %_Turnout, 2014



Map 10: Local R^2, Log Burglary vs. %_Turnout, 2010



Map 11: Local R^2, Log Burglary vs. %_Turnout, 2014

6.3.6 Summary of GWR Patterns

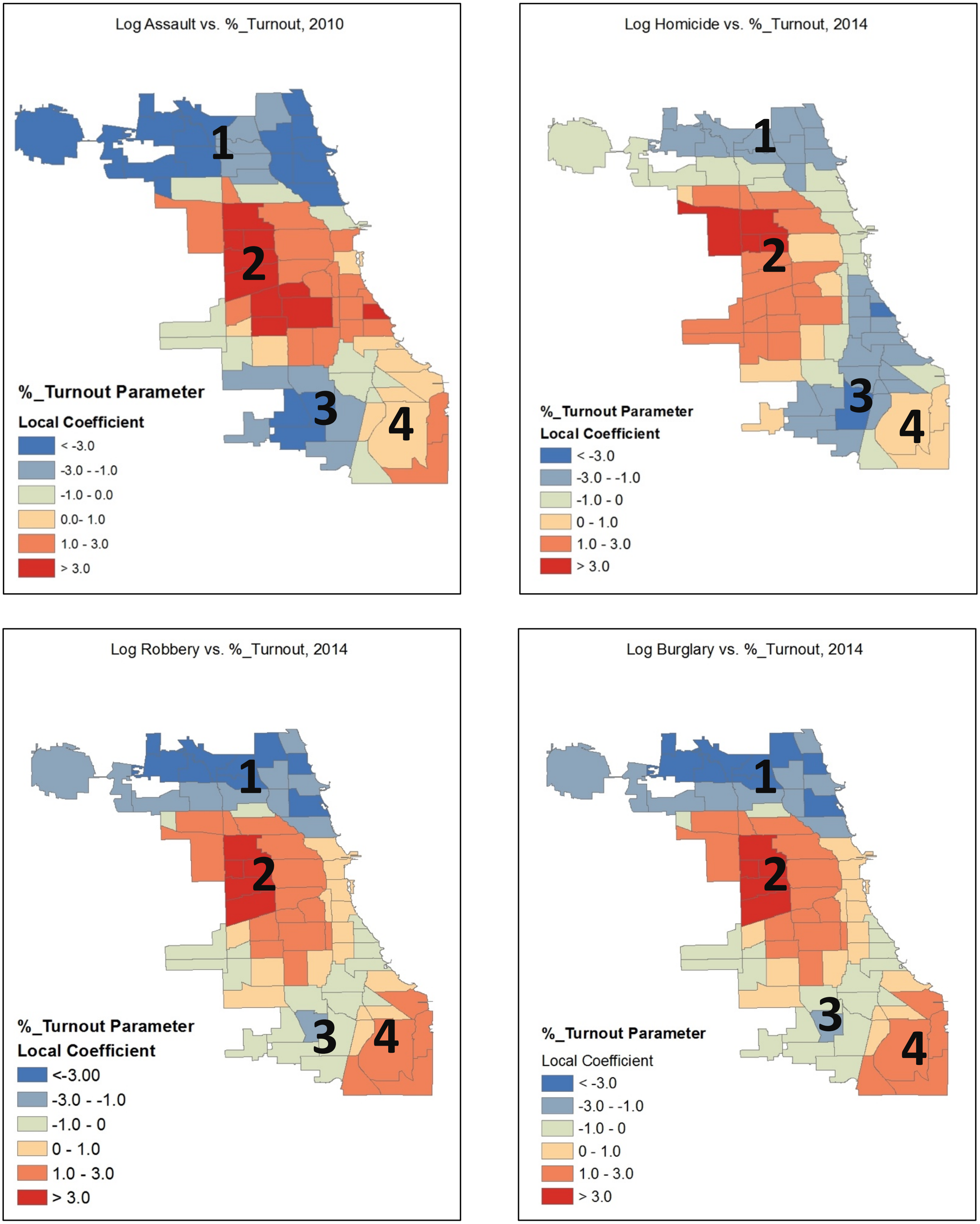
The GWR model does fairly well in explaining the relationship between different crime types and voter turnout. Moran's I was used to measure the degree of spatial autocorrelation present in the residuals. Across all crime types and years, the residuals were randomly distributed, indicating that the residuals are independent of one another and the errors are distributed randomly. The distribution of local R^2 values over different Chicago community areas reveals an interesting pattern: R^2 is highly variable over space, with some communities having an R^2 of over .80, and others with an R^2 of less than .20. The coefficients for the %_Turnout variable are also highly variable, with a negative sign in some community areas and a positive sign in others.

In comparing the R^2 values for the bivariate analysis of log crime vs. voter turnout, the model produces a reasonably high R^2 for all crime types, with rape having the lowest R^2 for both 2010 and 2014 (Table 3). This may be due to the fact that certain crime types such as homicide and rape, as Glaeser, Sacerdote, and Scheinkman (1996) point out, are less influenced by social conformity than other types of crime. Therefore, voter turnout, as a measure of social conformity, may explain less of the variation in rape compared to other types of crime.

	Log Homicide vs. %_Turnout	Log Rape vs. %_Turnout	Log Assault vs. %_Turnout	Log Robbery vs. %_Turnout	Log Burglary vs. %_Turnout
R^2 2010	.638	.473	.728	.782	.759
R^2 2014	.492	.491	.815	.799	.694

Table 3: R^2 for Log Crime Types vs. %_Turnout using GWR, years 2010 and 2014

Upon examination of the GWR maps depicting the relationship between voter turnout and crime, a distinct pattern can be observed that is consistent across multiple crime types and over time (years 2010 and 2014). Broadly speaking, there are four regions that consistently exhibit relatively high or low %_Turnout coefficients (Map 12). Region # 1, extending across Chicago's north side, and Region #3, on Chicago's southwest side, consistently have a negative coefficient for %_Turnout, meaning that in these areas, voter turnout is negatively associated with crime rate. Region #2, mostly centered around Chicago's west side, and Region # 4, on Chicago's southeast side, consistently have a high positive coefficient for %_Turnout compared to other areas of the city, meaning that the voter turnout and crime rates are positively associated in these areas.



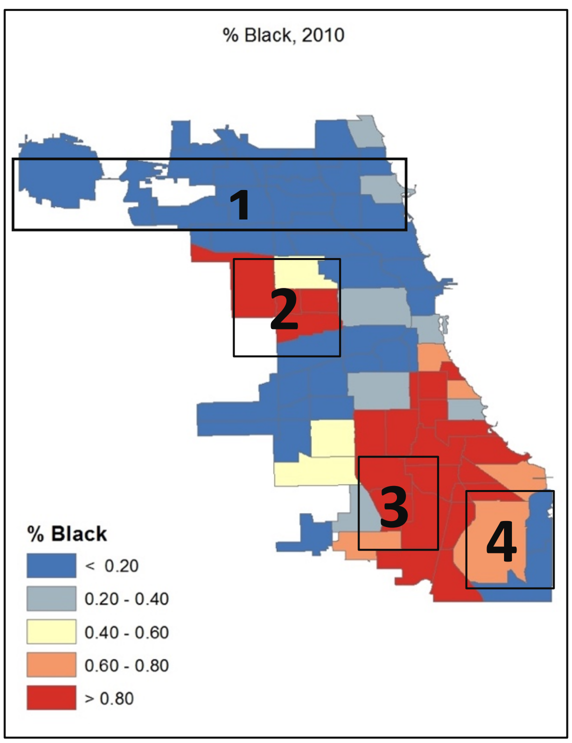
Map 12: Patterns of %_Turnout coefficient across multiple crime types and years

Given that the spatial pattern of the relationship between voter turnout and crime rate does not appear to be random, a descriptive approach can be taken in order to understand the spatial pattern. How can socioeconomic spatial patterns be used to explain the spatial patterns observed in the GWR analysis? The next section (6.3.7) addresses this question.

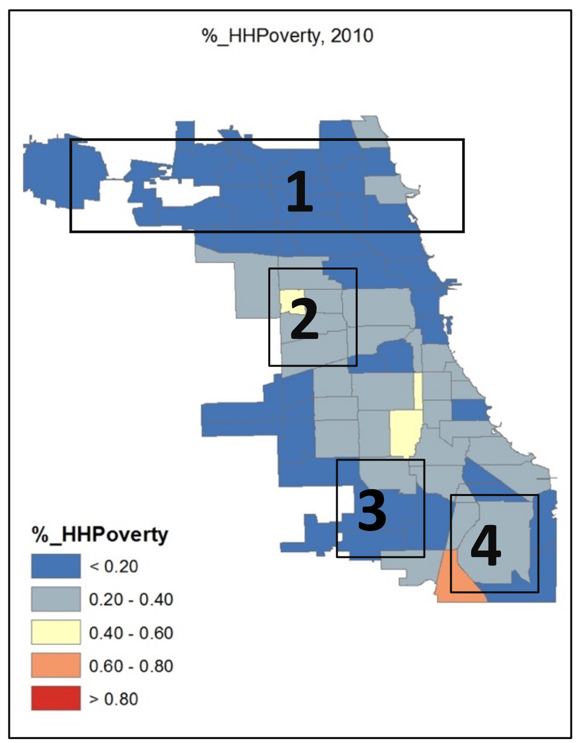
6.3.7 Spatial Distribution of Other Explanatory Variables

In order to explain the pattern observed for the relationship between voter turnout and crime rate, it is helpful to know whether the coefficient surface produced by GWR resembles the distribution of crime's explanatory variables, including race/ethnicity, income, poverty, and voter turnout. To compare the spatial pattern of the relationship between voting and crime with the spatial pattern of other explanatory variables, the spatial distribution of individual explanatory variables were mapped. The primary question is whether any of them exhibited the same spatial pattern as the voter turnout coefficient produced in GWR.

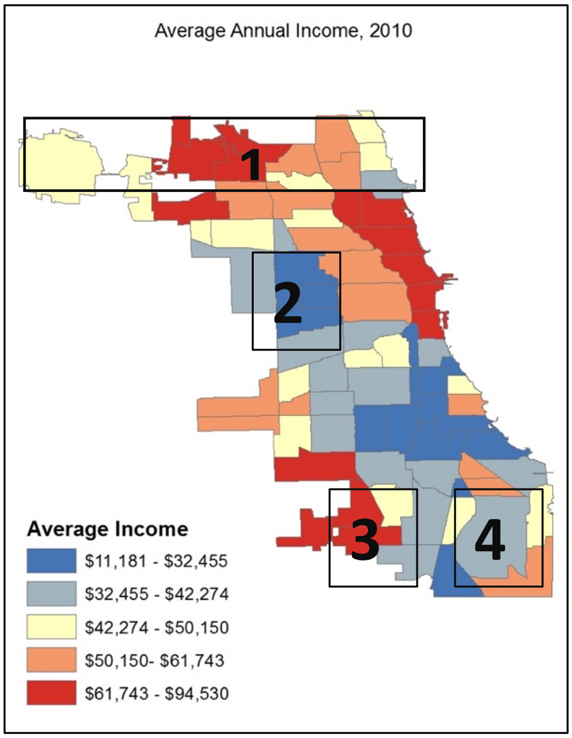
Of the six variables examined (%_Black, %_Latino, %_Minority, %_HHPoverty, Income, and %_Turnout), the spatial pattern of %_Black, %_HHPoverty, and Income appeared to mostly closely resemble the GWR voter turnout coefficient spatial pattern. The regions labeled on the %_Black, %_HHPoverty, and Income maps (Map 13, Map 14, and Map 15, respectively) correspond to the high and low %_Turnout coefficient regions depicted earlier in Map 43. Generally speaking, most high-coefficient community-areas in terms of %_Turnout were predominantly lower income and higher minority (high concentration of black, Latino, and/or Asian residents) while most low-coefficient community-areas in terms of %_Turnout were predominantly higher income and lower minority. An examination of the socio-economic and racial composition of the community-areas with the lowest and highest %_Turnout coefficients may help to explain the variation in the %_Turnout coefficient.



Map 13: %_Black

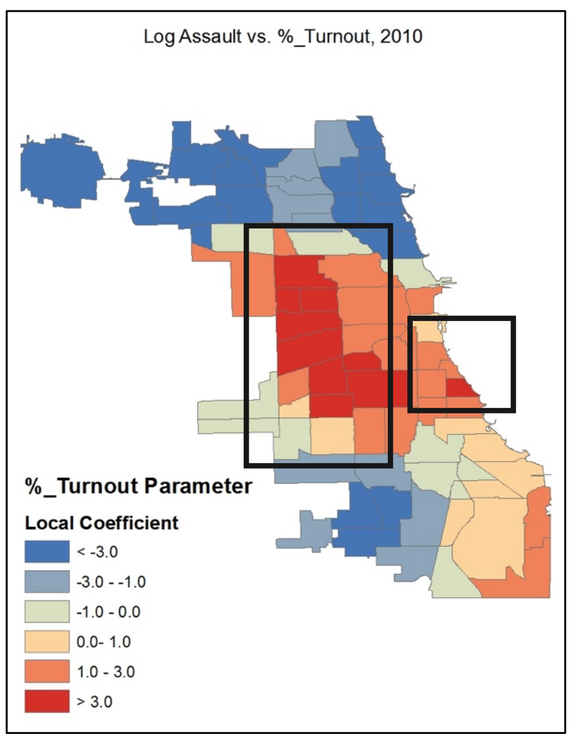


Map 14: %_HHPoverty



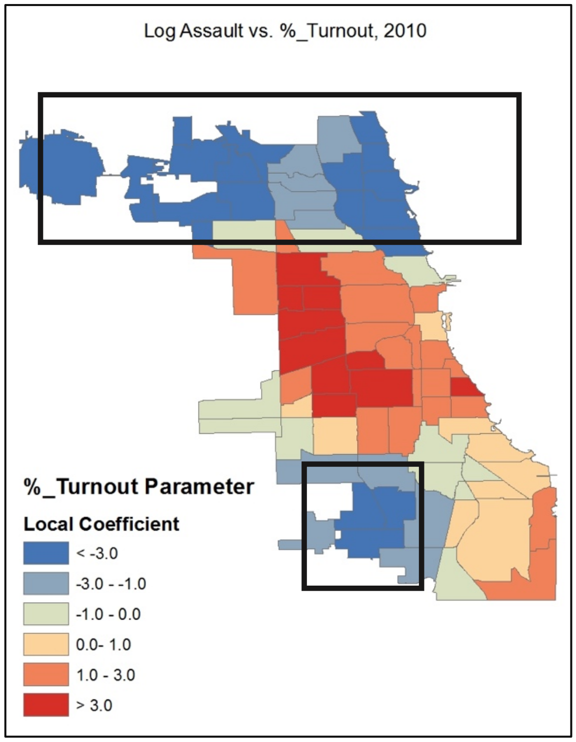
Map 15: Income

Using the example of log assault vs. %_Turnout for 2010 (Map 16), what are the characteristics of community-areas with a %_Turnout coefficient greater than 3.0? These communities are Brighton Park, East Garfield Park, Gage Park, Humboldt Park, Kenwood, McKinley Park, New City, North Lawndale, South Lawndale, and West Garfield Park. Generally speaking, these are all low-income, high minority, and low voter-turnout community-areas.



Map 16: Log Assault vs. %_Turnout, 2010
High %_Turnout coefficient community-areas

Similarly, using the example of log assault vs. %_Turnout for 2010 (Map 17), what are the characteristics of community-areas with a %_Turnout coefficient less than 3.0? These communities are Beverly, Dunning, Edgewater, Edison Park, Forest Glen, Jefferson Park, Lakeview, Lincoln Park, Lincoln Square, Montclare, Morgan Park, North Center, Norwood Park, O'Hare, Portage Park, Rogers Park, Uptown, and Washington Heights. Compared to the community-areas with high coefficients associated with %_Turnout, these community-community areas are relatively higher income, less minority, and exhibit higher voter turnout.



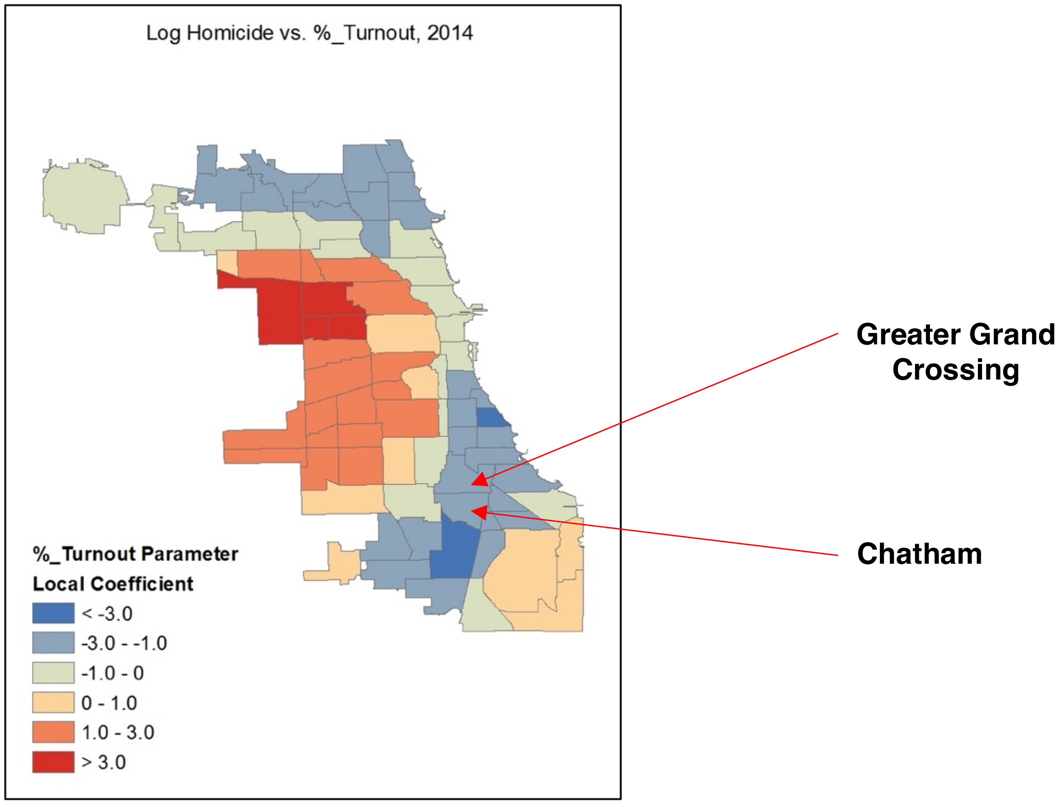
Map 17: Log Assault vs. %_Turnout, 2010
Low %_Turnout coefficient community-areas

Overall, there is a clear difference between community-areas with high coefficients associated with %_Turnout and low coefficients associated with %_Turnout. A side-by-side comparison of %_Turnout, Income, %_ HHPoverty, and %_Minority best illustrates this (Table 4).

	High %_Turnout Coefficient Community Areas (> .30)	Low %_Turnout Coefficient Community Areas (< -.30)
%_Turnout	.27	0.39
Income	$34,600	$64,450
% Households in Poverty	0.29	0.11
% Minority	0.93	0.38

Table 4: Comparing community-areas with high and low coefficients for %_Turnout

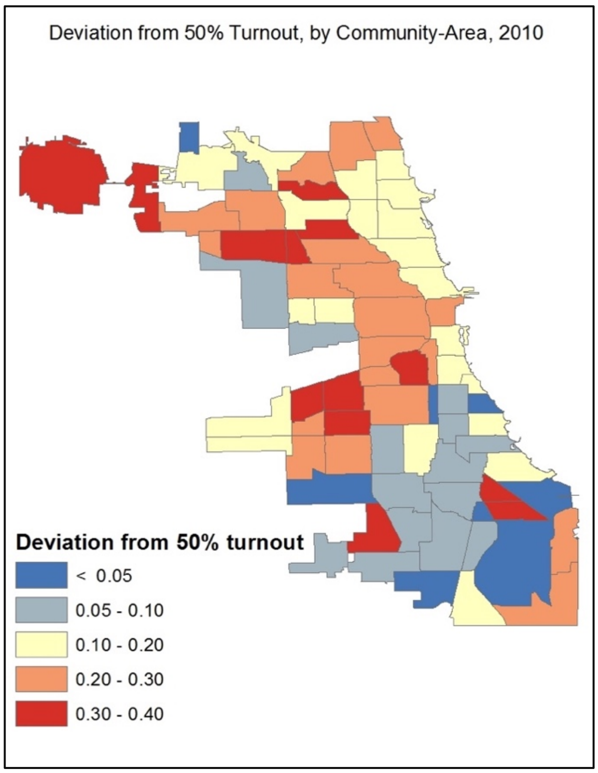
While it is clear that most community-areas with high coefficients for %_Turnout (regions #2 and #4) are lower income and higher minority, and most community-areas with low coefficients for %_Turnout (regions #1 and #3) are higher income and lower minority, not all lower income/higher minority community-areas fall into regions # 2 and #4, and not all higher income/lower minority community-areas fall into regions #1 and #4. For example, Greater Grand Crossing and Chatham on the south side of Chicago are nearly 100% black and low income, yet these community-areas don't fall into region #2 or #4 as would be expected. Rather, Greater Grand Crossing and Chatham both have low or even negative values for the %_Turnout coefficient in the homicide model (Map 18).



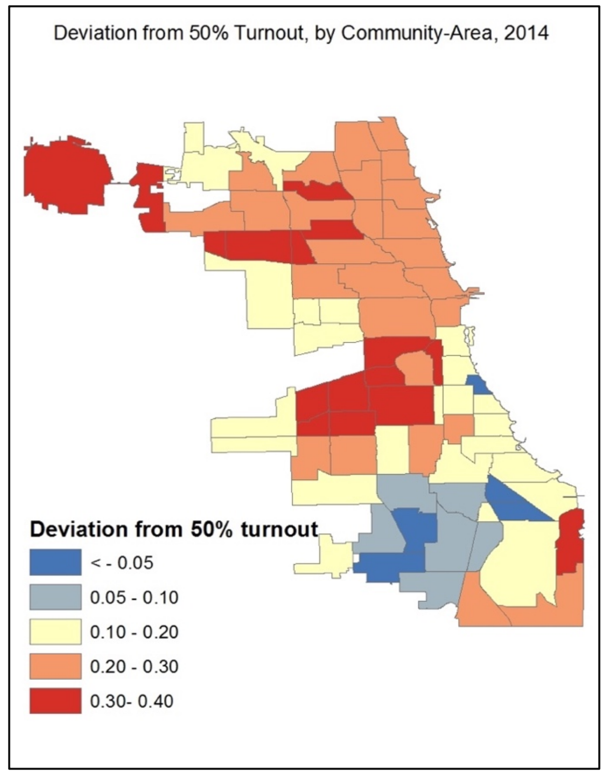
Map 18: Greater Grand Crossing and Chatham, south-side of Chicago

A possible conclusion is that percentage minority and socioeconomic status, as examined above, can only tell part of the story in terms of why some community-areas exhibit a positive relationship between voting and crime. As Coleman (2002) observed in a state and county-wide analysis, crime rate peaks where voter turnout is around 50% – the point at which there is non-conformity to the voting norm. A similar phenomenon may be occurring across Chicago community-areas. In certain community-areas undergoing transition, voter turnout might be approaching the 50% "entropic" mark. If residents of a particular community-area traditionally choose *not* to vote rather than vote, an increase in voting may nudge the community- area closer to the 50% turnout level, indicating a broad lack of consensus regarding the voting norm. A way to test this hypothesis would be to see whether there is a relationship between the voter turnout coefficients and deviations from 50% turnout. Theoretically, in community-areas with voter turnout hovering around 50%, an increase in voter turnout might push that community towards voting entropy (as described above) could subsequently lead to higher rates of crime. This could potentially explain how certain community-areas have positive associations between voting and crime.

Maps 19 – 20 show the spatial distribution of community-areas' deviation from 50% voter turnout. Interestingly, for 2010, three of the four community-areas with the highest coefficients associated with %_Turnout in the assault model (depicted in Map 16) deviated 20 percentage points or less from 50 % turnout. Still, a large number of community-areas with high %_Turnout coefficients have deviations between 30% and 40%, meaning that they are further away from the 50% turnout "entropy" mark but still have a positive association between voting and crime. Furthermore, the patch of community-areas on the southwest side of Chicago with voter turnout right around 50% has negative voter turnout coefficients (Map 17), indicating a negative relationship between voting and crime. Clearly, a community-area's proximity to 50% turnout cannot entirely explain the variation in the %_Turnout coefficients.



Map 19: Deviations from 50% turnout, 2010



Map 20: Deviations from 50% turnout, 2014

7. Conclusions

This research sheds light on the relationship between voter turnout and crime at the neighborhood-level. One of the primary goals of this work was to see whether or not the results from Coleman's (2002) study could be reproduced at a different scale of analysis – in this case – the neighborhood scale. In repeating Coleman's study at the Chicago community-area level, a few key insights are gained. First, at the community-area scale, there does not appear to be a strong association between voter turnout and crime rate. This is evidenced by the bivariate scatterplots of %_Turnout vs. various crime types, which show a weak association between the two variables. Furthermore, %_Turnout was a statistically significant predictor for only two of the five crime types used for this analysis. With the exception of assault and robbery in year 2010, %_Turnout was statistically insignificant as a predictor variable. Finally, this study compared crime models with and without the %_Turnout variable in order to assess how the addition of %_Turnout would impact R^2. For each crime type, when %_Turnout is added to Coleman's (2002) independent variables, the change in R^2 is minimal. It is clear that the addition of %_Turnout adds virtually no explanatory power to the model – a trend that is true across all five crime types.

This research also used stepwise regression to select optimal regression models for each crime type. For each crime type, Geographically Weighted Regression (GWR) was used to assess the spatial pattern of the relationship between %_Turnout and crime rate. For each crime type, local coefficients for %_Turnout, local R^2, and residuals were mapped in order to assess the spatial structure of the relationship between %_Turnout and crime rate. While there were minor variations amongst spatial patterns for voter turnout vs. various crime types, the GWR maps for all crime types depicted the same general spatial pattern and demonstrated that the relationship between voter turnout and crime rate is highly localized. The maps illustrate that there are clusters of Chicago community-areas that consistently exhibit a negative association between voter turnout and crime rate, and other community-areas where the association is consistently positive. Similarly, the local variation of R^2 seen in the GWR maps shows that voter turnout's ability to explain the variation in crime rate is highly variable over space.

Finally, this research attempts to explain what might underlie the observed patterns between voter turnout and crime rate. The spatial pattern of %_HHPoverty, %_Black, and Income bear close resemblance to the spatial pattern of voter turnout vs. crime rate, suggesting that the relationship between voting and crime may be partially mediated by socioeconomic and racial factors. Furthermore, examining the socioeconomic and racial composition of community-areas that have either strong negative or strong positive associations between voter turnout and crime confirms this. Community-areas with strong positive associations between voter turnout and crime rate are often low-income and high-minority, while community-areas with strong negative associations between voter turnout and crime are primarily high-income and low-minority.

When interpreting the GWR spatial patterns, an important caveat should be kept in mind. Knowing from the initial bivariate scatterplots illustrated in section 3.1 that %_Turnout and crime rate show a weak overall relationship, caution should be used when making assumptions about community-areas with high coefficients associated with %_Turnout and community-areas with negative coefficients associated with %_Turnout. It would be easy to say that in community-areas with high coefficients for %_Turnout, more voting means more crime, and in community-areas with negative coefficients for %_Turnout, less voting means more crime. However, voter turnout may encompass other factors and may be reflective of underlying characteristics in a particular neighborhood – an issue which this research only begins to address – and thus it is important not to make broad generalizations from the resultant GWR patterns.

Furthermore, GWR by nature focuses on local relationships. As such, GWR models runs the risk of being overly specified and too sensitive to local variation, and on the other hand, overly generalized and apt to miss important details. This research attempted to strike a proper balance between both extremes, but the risk of over-generalization and over-sensitivity with any GWR model nevertheless deserves mention here. Furthermore, this research does not attempt to explain how local factors present in various community-areas might influence the relationship between voting and crime, or how the analysis might be overly sensitive to local "noise", which further underscores the need to take a measured approach when interpreting the GRW spatial patterns.

Another important consideration is that any data aggregated to arbitrary spatial units is subject to the modifiable area unit problem (MAUP) – the idea that different units of analysis may produce different statistical results. The data in this study was aggregated up to the community-area, and it is possible that different results could be obtained if different spatial units of analysis were used. Furthermore, statistical results can vary depending on the overall scale at which the analysis is performed. Since this analysis was done only at the community-area scale, it is possible that alternative scales of analysis could produce more insights as to whether the observed relationships would hold with changes of scale.

Finally, the replicability of this research should be addressed – namely, whether or not the relationship between voting and crime observed in this research could be observed for cities other than Chicago. Though this research points to socioeconomic characteristics of community-areas and their level of voter turnout as two possible explanations of the observed relationship between voting and crime, no definitive explanation is given. Therefore, it is unclear whether socioeconomic factors, voter turnout, or both are responsible for the observed relationship. In a city with different socioeconomic characteristics and different levels of voter turnout, it is difficult to say whether the relationships in this study could be reproduced. Thus, future studies on these aspects would provide more insight on these relationships.

This original focus of this research was to better understand the relationship between voting and crime and the spatial structure of this relationship. In the process, new questions were brought to light – namely, what accounts for the unique spatial pattern of voter turnout and crime rate's relationship across Chicago community-areas? While this research begins to answer that question, the explanations offered are largely speculative. While this research addresses the overall relationship between voting and crime at the neighborhood level, and the spatial pattern of this relationship, understanding the underlying causes that impact how voter behavior and crime interact over space is deserving of further research.

References

[1] Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, *58*(6), 1015–1026.

[2] Cialdini, R. B., & Trost, M. R. (1998). Social Influence: Social Norms, Conformity, and Compliance. *The Handbook of Social Psychology*, 1(4), 151-192.

[3] Coleman, S. (2002). a Test for the Effect of Conformity on Crime Rates Using Voter Turnout. *The Sociological Quarterly*, *43*(2), 257–276.

[4] Coleman, S. (2004). The effect of social conformity on collective voting behavior. *Political Analysis*, *12*(1), 76–96.

[5] Foster, C. B. (2016). The Performance of Rational Voter Models in Recent Presidential Elections. *The American Political Science Review*, 78(3), 678-690.

[6] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. West Sussex, UK: John Wiley.

[7] Gerber, A. S. (2004). Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You, 178–191.

[8] Gerber, A. S., Green, D. P., Larimer, C. W., & Iowa, N. (2008). Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment, *102*(1), 33–48.

[9] Knack, S. Civic Norms, Social Sanctions and Voting Turnout. *Rationality and Society*, 4(2), 133-156.

[10] Land, K. C., McCall, P. L., & Cohen, L. E. (1990). Structural Covariates of Homicide Rates: Are There Any Invariances Across Time and Social Space? *American Journal of Sociology*, *95*(4), 922.

[11] Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence. *Criminology*, *39*(3), 517–560.

[12] Riker, W. H. (1968). A Theory of the Calculus of Voting. *The American Political Science Review*, 62(1), 25-42

[13] Rosenfeld, R., Messner, S. F., & Baumer, E. P. (2001). Social Capital and Homicide. *Social Forces*, *80*(1), 283–310.

[14] Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science, New Series*, *277*(5328), 918–924.

[15] Uggen, C., & Manza, J. (2004). Voting and Subsequent Crime and Arrest: Evidence from a Community Sample. *Columbia Human Rights Law Review*, *36*(193), 193–215.