A GEOSPATIAL ASSESSMENT OF DOMESTIC VIOLENCE

Author:
Lauren Holt

Faculty Sponsor:
David Holleran,
Department of Criminology

ABSTRACT
This paper conducts a geospatial assessment of domestic violence in Lawrence Township, NJ in order to better understand the nature of the domestic violence crimes in that area. This entails using dispatch data from the Lawrence Township Police Department and applying statistical procedures on the risk factors associated with domestic violence. The risk factors under consideration in this paper, as identified by prior research, are age, gender, renter status, prior instances of domestic violence, and noise complaint calls. The results of the research suggest that not all risk factors as identified by prior research are associated with domestic violence in Lawrence Township. Furthermore the results indicate that the risk factors explored in this paper have a minimal effect individually, but may have a more meaningful explanation when used together. As a result, other factors should be explored because they may have a more meaningful relationship with domestic violence in Lawrence Township, NJ than those examined in this paper. Thus continued research is encouraged.

INTRODUCTION
Domestic violence is a distinctive and prevalent crime in the United States. In 2010 there were 74,244 domestic violence offenses reported to the police in New Jersey (State of New Jersey Department of Law and Public Safety, 2010). The majority of offenses in 2010 were assaults (42 percent) and harassments (44 percent), and these percentages have remained relatively stable over the past five years (State of New Jersey Department of Law and Public Safety, 2010). Domestic violence victims often continue to face victimization and certain offenders will repetitively victimize a variety of people (Klein, 2009). Domestic violence incidents also tend to be greatly underreported within the criminal justice system (Mannon, 1997). Although these crimes tend to be underreported, they still consume a large proportion of police time and resources, accounting from anywhere to 15 to more than 40 percent of calls (Klein, 2009; Hendricks, 1991). Thus it is imperative to recognize the large impact domestic violence crimes play and push to better understand this unique crime.

LITERATURE REVIEW
Domestic violence studies have identified a multitude of risk factors associated with domestic violence events, of which many have examined variables concerning both victims and offenders. Risk factors include, but are not limited to, age, gender, economic status, employment status, and substance abuse.

Research has shown that people of all ages are at risk for domestic violence. However certain age groups were more likely to be associated with domestic violence than others. Previous studies demonstrated that the majority of perpetrators were between the ages of 18 and 35 (Klein, 2009; Belknap, Graham, Hartman, Lippen, Allen, & Sutherland, 2000; Buzawa, Hotaling, Klein, & Byrnes, 1999). Furthermore, in a survey involving 218 battered women at the San Francisco General Hospital Medical Center, 74 percent of the victims were between the ages of 16 and 35, and 58 percent of the male batterers were between the ages of 16 and 35 (Berrios & Grady, 1991). Victims reported that offenders tended to be about the same age as them (Catalano, 2007). In general, domestic violence crimes have decreased as couples have aged (Straus et al., 1980; Riggs, Caulfield, and Street, 2000). For example, research has found that domestic violence is primarily concentrated in couples younger than 30 years old (Straus et al., 1980; Wilt and Olson, 1996). Other research has determined a narrower age-group range; males age 20 to 24 are at the greatest risk of being perpetrators for nonfatal intimate partner violence (Catalano, 2007). In general, younger populations tend to have more occurrences of domestic violence victimization and/or perpetration than older populations.
Perpetrators of domestic violence have been both male and female. However the majority of studies suggested that males are overwhelmingly perpetrators of domestic violence (Klein, 2009). Wordses found that in Berkeley, California, 84 percent of the reported suspects were male (2000). Conversely other studies have found that males and females offend equally (Klein, 2009; Straus, Gelles, & Steinmetz, 1980). For example, numerous studies have found that a sizeable amount of women hit or beat their husbands (Kimmel, 2002). However, when the severity of the offense was taken into account, male offenders broke criminal and civil laws more frequently than female offenders did (Macmillan & Krutttschnitt, 2004; Klein, 2009). Additionally males were more likely to inflict more serious injuries than women were (Kimmel, 2002). Lastly, as much as 75 percent of violence committed by women was done in self-defense (Kimmel, 2002). Overall, previous studies have shown that the majority of males were perpetrators, but that females were aggressors as well. Correspondingly victims of domestic violence were predominantly females. Some studies suggested there is not a gender disparity in victimization and that males were victimized as frequently as females (Klein, 2009). However the National Crime Victimization Survey found that women were ten times more likely to be victims of domestic violence than men (Zawitz 1994; Friday, Lord, Exum, and Hartman, 2006). Furthermore according to the 2010 National Crime Victimization Survey, there were 544,730 instances of intimate partner violence in which a female was the victim and 125,120 instances in which a male was the victim (Truman, 2011). Thus in most previous cases of domestic violence, males were perpetrators and females were victims.

Domestic violence has frequent instances of repeat victimizations and repeat offending. In a study of battered women from the San Francisco area, 86 percent of the victims reported at least one previous occurrence of abuse, of which 40 percent required medical attention (Berrios & Grady, 1991). Additionally a study in Rhode Island found that 38 percent of offenders were arrested for a new domestic violence offense after being charged with a misdemeanor domestic violence offense (Klein, 2009; Klein, Wilson, Crowe, and DeMichele, 2005). Furthermore, Friday, Lord, Exum, and Hartman found that the best predictor of future domestic violence incidents was a prior record of domestic violence (2006). Offenders have also been known to abuse multiple victims, including other intimate partners or family members. Thus repeat occurrences were not limited to the same victim. For example, in a probation study, it was found that in a one-year period 28 percent of the offenders were arrested for a new domestic violence offense against a new victim (Klein, 2009; Klein, Wilson, Crowe, and DeMichele, 2005). Although repeat victimization and offending does not always occur, it nevertheless has been prevalent in certain domestic violence crimes.

Lastly, the type of dwelling people lived in was correlated with domestic violence. People who lived in renter-occupied housing units were at a greater risk of victimization than people who lived in owner-occupied housing units. Average annual domestic violence rates were considerably higher for those living in rental housing than other types of housing (Catalano, 2007). For both males and females, the rates of victimization were more than three times the rate of those who lived in an owned house (Catalano, 2007).

**METHODOLOGY**

Lawrence Township is a suburban community located in Mercer County, New Jersey, and it had a population of 33,472 as of the 2010 United States Census (United States Census Bureau, 2011). It is located most notably between Princeton Township to the northeast (population 16,265) and Trenton to the southwest (population 84,913), and it contiguously borders six townships (United States Census Bureau, 2011). The median household income amongst people in all occupied housing units as of 2010 was $86,870 (United States Census Bureau, 2010). People residing in owner-occupied housing units had a median income of $106,603 whereas people residing in rental-occupied housing units had a median income of $47,974 (United States Census Bureau, 2010). Although the median household income in Lawrence Township was above the national and the Northeast average, there was considerable variability in household income when considering the type of residential unit (DeNavas-Walt, Proctor, & Smith, 2012). The majority of residents in Lawrence Township were White (70.36 percent), but large percentages of Asians (13.22 percent) and Blacks or African Americans (12.87 percent) also resided there (United States Census Bureau, 2010). As of 2012, multiple motor vehicle arterial paths converged within Lawrence Township, including Interstate-95, Interstate-295, U.S. Route 1, and U.S. Route 206, among
other major county roads. Furthermore there were six different New Jersey Transit bus routes in 2012 that traveled between Lawrence Township and neighboring municipalities. Thus Lawrence Township experienced a great deal of vehicular movement despite being a suburban community.

The data used in this study were obtained from the Lawrence Township Police Department (LTPD) 2010 dispatch file. The data were first imported and modified, where necessary, in SAS/STAT 9.3. Modifications entailed removing sensitive information and creating fields necessary for geocoding domestic violence events from existing fields. For example, an address field was created from the ‘street number’, ‘street name’, and ‘street suffix’ fields. The data were then examined in SAS/STAT 9.3 utilizing various statistical procedures. In 2010 the LTPD responded to 179 calls of alleged domestic violence. The mean time spent by responding officers on a domestic violence call was 242 minutes, with a minimum time of 34 minutes and a maximum time of 616 minutes. Domestic violence calls were nearly uniformly distributed each month with approximately 13 to 20 calls each month. However, December and January had fewer calls than usual, with only 10 and 6 respectively. In terms of domestic violence calls per quarter, the third and fourth quarters exhibited more calls than the first and second quarter. More calls were received during the PM hours than the AM hours; 68 percent of all calls occurred during the PM hours.

A shapefile of Mercer County, New Jersey, was obtained from the U.S. Census, Topologically Integrated Geographic Encoding and Referencing (TIGER) files (United States Census Bureau, Geography Division, 2010). Three different shapefiles were obtained which feature the Mercer County street, tract, and block levels. The three different shapefiles of Mercer County were imported into ArcGIS Version 10.1. None of the geospatial analyses involved areas outside of Lawrence Township; thus, all areas outside of Lawrence Township were removed using the clip feature in the ArcMAP toolbox so that only Lawrence Township appeared in the final shapefile images. The shapefiles featuring the Lawrence Township block, street, and tract levels were then changed from a geographical coordinate system to a projected coordinate system: Universal Transverse Mercator Zone 18 North.

The Lawrence Township block level map would prove to be the most integral in this research project for a variety of reasons. The block levels were Census designated geographic areas which approximated a city-block. In a city, this area was ideal, but in suburban locations it was less applicable. Nevertheless, block level maps were the smallest geographic areas which the U.S. Census releases demographic information for. Thus when examining the various geographies within a singular township, block level maps were appropriate because they conveyed more precise information than a tract level map did. For example, there were 571 census designated blocks in Lawrence Township but only 5 census designated tracts. Thus the individual blocks (n=571) constituted the unit of analysis in this study.

**DEPENDENT VARIABLE**
The dependent variable was the number of domestic violence calls per block from the 2010 LTPD dispatch data. The LTPD dispatch data did not contain the longitude and latitude geographic positioning for each domestic violence call. The data did contain the address for each event, which would make it possible to determine the global position of each domestic violence call. It was necessary to utilize the geocoding tool in ArcGIS 10.1, which provided the approximate global position for each event based on the given address. The majority of the data events were geocoded using the geocoding tool in ArcGIS 10.1, but the remaining, unidentified locations were geocoded using Abby is Queen, a geocoding website (http://www.abbyisqueen.com/). A comprehensive map was rendered through this process which featured the street, block, and tract layers of Lawrence Township, as well as the domestic violence calls as X,Y coordinate points on the map. The 2010 domestic violence call points were joined with the Lawrence Township block level map layer. The joined map layer provided a count of the domestic violence calls for each U.S. Census designated block.

**TEST FOR SPATIAL RANDOMNESS**
A key starting point with the geospatial analysis was to determine whether the domestic violence calls were significantly clustered or randomly clustered in particular locations of Lawrence Township. Using the spatial analyst extension toolbox in ArcGIS 10.1, the plotted domestic violence call point data were
tested for spatial randomness. The null hypothesis for spatial randomness was that the distribution of points is randomly distributed. The Global Moran’s I tool was used to determine spatial randomness. A fixed distance band was used for conceptualization of spatial randomness, and Euclidean distance was selected for the distance method. Nine different distance bands were tested to determine which distance band would be appropriate for use in a hotspot analysis.

Distance bands equal to or less than 3500 ft. did not result in enough neighbors and thereby resulted in an error. 4 4000 ft. was the smallest distance in which the number of neighbors was large enough so each domestic violence call had at least one neighbor, and thus the test was able to run without failure. Thus it was determined that 4000 ft. was the optimal distance in analyzing this particular set of domestic violence events. The Global Moran’s I tool test determined that the z-score for the distance band of 4000 ft. was .194065, indicating that the null hypothesis could not be rejected and that the domestic violence call points are randomly distributed. In other words, it could be assumed that the spatial processes were random. The points then were not more or less spatially clustered than would be expected if the spatial processes were random. See Table 1 for the Z Score of all distance bands explored.

Table 1. Test for Spatial Randomness. At 4000 ft. the Z Score was .194065, which was not statistically significant.

<table>
<thead>
<tr>
<th>Distance (x)</th>
<th>1500 ft.</th>
<th>2000 ft.</th>
<th>2500 ft.</th>
<th>3000 ft.</th>
<th>4000 ft.</th>
<th>4500 ft.</th>
<th>5000 ft.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z Score (y)</td>
<td>1.776961</td>
<td>1.030137</td>
<td>1.492389</td>
<td>1.95447</td>
<td>.194065</td>
<td>.323590</td>
<td>-.1301024</td>
</tr>
</tbody>
</table>

A hot spot analysis was then conducted utilizing the Hot Spot Analysis (Getis-Ord Gi*) tool in ArcGIS 10.1. Hot spot analyses display where there are high or low clusters, which are areas with either high values or low value surrounded by similarly high or low values. The domestic violence call counts were used as the input feature. The hot spots (GiZScore > 1.65 Std. Dev.) and cool spots (GiZScore < -1.65 Std. Dev.) were accordingly determined. The results are displayed in Figure 1, showing the hot and cool spots as calculated by the hot spot analysis. The hot spot analysis shows where the highest spatial clustering of domestic violence events were for the 2010 data based on the count of domestic violence calls. All of the domestic violence hotspots were clustered in the southern half of Lawrence Township.

A kernel density image was also created using the Spatial Analyst extension in ArcMap 10.1, which utilized the same radius (4000 ft.) as the hot spot image. The density image (Figure 2) shows where the greatest prominence of domestic violence calls were in 2010, with the highest concentration in the red zones.
Figure 1 (on left) and Figure 2 (on right). The legend values in Figure 1 display the statistical significance of the spatial clustering within a 4000 ft. radius. The blue blocks represent cold spots and the red blocks represent hot spots. White blocks indicate areas that are not statistically significant.

The hot spot analysis and the kernel density image conveyed that there are indeed areas within Lawrence Township that had higher spatial clustering than would be assumed by pure randomness. Thus it became important to try and explain why these hot spots are concentrated in particular areas. Prior research has attempted to explain certain risk factors associated with domestic violence. I attempted to examine the factors explored in the Literature Review section of this paper to determine if they are relevant in understanding the domestic violence landscape in Lawrence Township, NJ.

INDEPENDENT VARIABLES
The next phase of the analysis entailed an examination of the factors that correlate with the domestic violence events in Lawrence Township. All of the independent variables under consideration had to be aggregated to the block level. In other words, a count per block for each variable had to be calculated. The independent variables under consideration were the percentage of females age 20 to 29, the percentage of males age 20 to 29, the percentage of renters, the noise abated calls from the 2010 dispatch data, and the domestic violence calls from the previous year’s (2009) dispatch data. The percentage of males age 20 to 29, the percentage of females age 20 to 29, and the percentage of renters were selected based upon the prior research in the Literature Review section of the paper. The 2009 domestic violence count events were used as a control to determine the variation of the domestic violence landscape over
L. HOLT: A GEOSPATIAL ASSESSMENT OF DOMESTIC VIOLENCE

the past year. Prior research suggests that domestic violence crimes are often repetitive and not isolated occurrences. Thus the 2009 domestic violence count was used as a control within the independent variables to determine how much change there was in 2009 and 2010 alleged domestic violence count. Lastly, the noise abated calls from 2010 were used as an independent variable to see if noise complaint calls were associated with domestic violence occurrences.

The independent variable information sets were obtained from a variety of sources. Two of the independent variables under consideration, the 2009 domestic violence calls and the 2010 noise abated calls, were taken directly from the 2009 and 2010 LTPD dispatch data. These independent variables were geocoded in the same exact way that the 2010 domestic violence calls were, as indicated under the dependent variable section. They were given an X,Y coordinate position and then aggregated to the block level.

Three of the independent variables were obtained as block level data tables from the U.S. Census American Fact Finder (factfinder2.census.gov/). The data tables selected from the American Fact Finder for this study were ‘age by sex’ and ‘tenure’. The data tables for each variable were imported into SAS 9.3 in order to make them compatible with the Lawrence Township shapefiles in ArcGIS 10.1. In order for the independent variable data tables to be joined with the Lawrence Township shapefile, they had to have an identical field in common in their data tables. The shapefile of Lawrence Township contained the identification field ‘GEOID’ and the independent variable data tables obtained from the American Fact Finder similarly had a field ‘Id’, which contained the identification code from the GEOID variable. The SUBSTR command was used in SAS 9.3 to create a new field, GEOID, from a portion of the Id variable.

The independent data table layers were then able to be combined with the Lawrence Township block level domestic violence call join map. All of the data tables were combined to form one large geodatabase of risk factors. The joined map contained a count of all the independent variables and dependent variables per each block in Lawrence Township. This joined map made it possible to conduct geospatial and statistical analyses in ArcGIS 10.1 and SAS 9.3. The map table was exported from ArcGIS and imported into SAS 9.3.

Several of the independent variables were converted from a raw count to a percentage in SAS 9.3. The ‘tenure’ variable provided the count of the total number of home-owners, renters, and population in housing units in Lawrence Township. This count was then converted to a rate, the number of renters divided by the population in housing units and multiplied by 100 for a percentage. The percentage of renters, rather than the total count for renters, was used as the final independent variable.

The variable ‘age by sex’ was also converted from a raw count to a percentage. The original data table provided the number of people, by their sex, within a certain age group. For example, how many females age 22 to 24 there were per block. The different variables were combined to form a large age range variable, males age 20 to 29 and females age 20 to 29. This involved adding together the variables M20+M21+M22_to_24+M25_to_29 and F20+F21+F22_to_24+F25_to29. The variables males age 20-29 and females age 20-29 were each separately divided by the total population and multiplied by 100. The percentage of males age 20-29 and females 20-29 were used as the final independent variables.

The other independent variables did not need to be converted, as the raw count would be appropriate. All independent and dependent variables were quantitative in nature and based on a ratio scale. Based on previous literature, the risk factors should be statically significantly correlated with an increased count of domestic violence events. The unit of analysis is the U.S. Census block level. Lawrence Township was broken down into 571 different block areas, making it an appropriate size for conducting analyses.

RESULTS

After the geodatabase of risk factors was modified in SAS 9.3, the data were ready to be statistically analyzed. The dispatch data for the 2009 and 2010 domestic violence events displayed that the distribution of events was right-skewed. There was a preponderance of zeros, meaning that the majority of U.S. Census blocks in Lawrence Township had no occurrences of domestic violence. Given that the dependent variable constituted a count of integer data, count regression models such as Poisson regression and negative binomial regression were explored to examine the conditional distribution of
domestic violence counts and the independent variables in the analysis. Because the data exhibited what is known in the statistical literature as “overdispersion”, a negative binomial regression was used in place of Poisson regression.

The negative binomial regression procedure was performed in which the dependent variable was set to the count of 2010 domestic violence call count and the independent variables were set as the percentage of renters, the percentage of males age 20-29, the percentage of females age 20-29 rate, the 2009 domestic violence call count, and the 2010 noise abated call count. There were 205 blocks in Lawrence Township that were designated as non-residential according to the 2010 U.S. Census. Out of the 205 non-residential blocks, alleged domestic violence events occurred in only 5 of the blocks, representing 14 of the total 179 events (7 percent). As a result, non-residential blocks were omitted from the statistical analyses.

The negative binomial regression estimates of domestic violence events per block are presented in Table 1. Three of the variables were statistically significant whereas two of the variables were not. The 2009 domestic violence count, the 2010 noise abated count, and the percentage of renters were all statistically significant (i.e., p<.05). Thus, areas with a high concentration of renters, previous domestic violence occurrences, and/or noise abated calls were associated with a higher concentration of domestic violence events. Conversely, there was no evidence that areas with a higher concentration of males age 20-29 and/or females age 20-29 was associated with a higher concentration of domestic violence events when the variables in Table 1 were considered. Based on the results, the location of the 2009 domestic violence events did not change much in comparison to the 2010 domestic violence events, and thus those areas could continue to have a problem with domestic violence crimes.

Table 1

| Variables                        | DF | Parameter Estimate | Standard Error | t Value | Approx Pr > |t| |
|----------------------------------|----|--------------------|----------------|---------|-------------|------|
| Intercept                        | 1  | -1.729299          | 0.179965       | -9.61   | <.0001      |
| Count domestic violence calls 2009 | 1  | 0.398833           | 0.132004       | 3.02    | 0.0025      |
| Count noise abated calls 2010    | 1  | 0.271787           | 0.068669       | 3.96    | <.0001      |
| Females 20 to 29 percentage      | 1  | 0.035277           | 0.024442       | 1.44    | 0.1489      |
| Males 20 to 29 percentage        | 1  | 0.003742           | 0.021481       | 0.17    | 0.8617      |
| Rent percentage                  | 1  | 0.007364           | 0.003626       | 2.03    | 0.0422      |

The predicted domestic violence count was calculated for every block utilizing the negative binomial regression prediction equation. The predicted values for each block were exported from SAS 9.3 and imported in ArcGIS 10.1. The predicted values were symbolized for each block and are displayed in Figure 3. Figure 4 represents the actual amount of domestic violence calls for each block based on the 2010 dispatch data. The predicted equation map and the actual count map contain many blocks that display similar results. Although the map images appear to be quite similar, the numeric scales are not identical. The predication equation map has a larger numeric range than the actual count map does. Thus, in predicting an actual count of domestic violence crimes, the predication equation is off. However it is still accurate in pinpointing which areas have high or low counts in comparison to other areas. Therefore the predication equation used in this study may still be useful in understanding the geography of domestic violence in Lawrence Township.
Figure 3 (on left) and Figure 4 (on right). Figure 3 displays the predicated count based on the negative binomial regression equation and Figure 4 displays the actual count.

CONCLUSION
A geospatial assessment of domestic violence in Lawrence Township, NJ was conducted in order to try and better understand the landscape of alleged domestic violence crimes. The independent variables under consideration are the percentage of renters, the percentage of males age 20 to 29, the percentage of females age 20 to 29, the count of domestic violence calls from the previous year (2009), and the count of noise abated calls from 2010. The results signify that not all of the risk factors utilized in this study are correlated with domestic violence events in Lawrence Township. Furthermore, the risk factors which do have a statistically significant t-value do not have overwhelmingly large parameter estimates, and so the effect of each independent variable is not very powerful. The variable domestic disturbance also may have had a particularly strong correlation due to a preponderance of zeros, and thus must be interpreted cautiously. It also may have contributed to the particularly large predicted count (51) in one of the blocks in Figure 3. Although many of the independent variables were not statistically significant, the predication equation as a whole was able to create similar results in comparison to the actual results. Thus there may be some merit in observing the risk factors as a whole rather than individually.

The study relied upon reported domestic violence calls only. Although many crimes go unreported, domestic violence events in particular are vastly underreported (Mannon, 1997). Reported incidents of domestic violence constitute only a portion of the whole picture. It is important then to keep
in mind that perhaps certain demographics have a greater tendency to not report domestic violence crime than others, which has a possibility to throw off certain inferences made. Additionally the data represents the entire data set of alleged calls. No sampling methods were used, and inferences from the negative binomial regression must be made cautiously. As a result findings from this study are limited to Lawrence Township only; it would be inappropriate to infer characteristics from this study onto the entire population of townships in Mercer County, New Jersey, or the United States. Unfortunately the U.S. Census releases minimal information for individual block. Thus, this research project was limited in the independent variables it could examine. In order to better understand domestic violence, a variety of other risk factor should be explored as well. For example, race, income, and marital status could be considered as well.

The results do provide some relevant considerations for Lawrence Township. Some of the risk factors studied here were not statistically significant, but as a whole the independent variables may have some merit. Additionally there is evidence that suggests other risk factors should be examined in order to better explain and understand domestic violence in Lawrence Township. There are certainly pockets within Lawrence Township that exhibit more instances of alleged domestic violence than others, and so continued research should be encouraged.

REFERENCES


- 9 -


NOTES

1. The three fields were concatenated in a SAS/STAT data step using a concatenation function for string data (i.e., CATX).

2. The calls were deemed alleged because they were classified as domestic violence during the police contact only and are thus predisposition. The status of the calls and how they have proceeded through the criminal justice system is not accounted for with the Lawrence Township Police Department data.

3. Shapefiles were obtained for the Mercer County street, tract, and block levels. The block level shapefile was the only one utilized for geospatial analyses. The street and tract-level shapefiles were used as references.

4. The shapefile was converted from a geographic coordinate system to a projected coordinate system in order to utilize ArcGIS applications. The UTM coordinate system is particularly advantageous for areas the size of a U.S. state or smaller, and thus it was selected (Esri, 2009).
5. Neighbors in this case entail other instances of domestic violence calls. Due to the skewed nature of the data it was imperative to obtain a distance band in which each domestic violence call point had at least one other domestic violence call point within its radius (4000 ft. in this case).

6. The ‘Id’ field contained more characters than the ‘GEOID’ field, and so only part of the ‘Id’ field had to be used.