The Significance of Vacant Lots with Respect to Crime Rate, Time of Crime, and Type of Crime

Nicholas Bergh, Kaley Hanrahan, Andrew Pomykalski, Muyang Sun

Abstract

Crime and violence remain one of the most significant public issues in the United States. It has been found that there is strong relationship between crime rates and vacant properties. We tested this hypothesis and took it to a deeper level by classifying crimes by time of day, day of week, and type of crime and comparing their respective distributions near vacant properties versus non-vacant properties. We feel that the more the police force know about where and when crimes are likely to occur, the easier it would be for them to get the needed resource for their beats and prepare for crime in order to increase their own safety. By taking a deeper investigation into the Chicago crime data, we hope to provide a methodology that is applicable to informing law enforcement across all cities and regions in America.

1. Introduction

Concern about crime and violence in the U.S. is at a 15-year high, with 53% of Americans stating they worry a 'great deal' about crime (Davis, 2016). Law enforcement practices are being called into question, as members of leading law enforcement organizations advocate for a data-driven, modern approach to crime. States that have adopted more targeted approaches, targeting violent crime in particular, have seen crime and incarceration rates fall (Holden & Serpas, 2016).

Crime data in major cities is the most researched due to the wealth of data that is publicly available. That being said, crime in some cities including Memphis and St. Louis has risen over the past three years. Also, cities have violent crime rates well above the national average of 365.5 per 100,000 inhabitants with Detroit being the leader with 1,989 violent crimes per 100,000 inhabitants (Fisher, 2015).

Plenty of factors that may cause violence have been well investigated (Raphael & Winter-Ebmer, 2001; Sampson et al., 1997; J. Blau & P. Blau, 1982). There has been research done before that correlates vacant buildings and crime that have given very good results with strong correlation (Branas et. al, 2012; Spelman 1993). Decaying and abandoned neighborhoods are often accompanied with high crime rates, because vacant lots and buildings potentially offer

refuge to criminal and other illegal activities, and they also visibly symbolize that a neighborhood has deteriorated (Spelman, 1993; Kraut, 1999).

In this paper we examine the topic of vacant lots and crime. Based on existing literature, we believe that rate of crime is higher on vacant lots than non-vacant lots. In order to develop a deeper understanding of the effect of vacant lots on crime patterns, we also wanted to examine how the attributes of crime differ given location type. We examined the distribution of crime types and time of crime occurrences for vacant and non-vacant lots. We looked at time of day, the day of week, and the season in which a crime occurs. If crime rates are higher on vacant plots, an understanding of when the crimes occur as well as the types of crime that occur is of utmost importance.

2. Data

We focused on three datasets for the entirety of the analysis. These three datasets were provided to us by the City of Chicago's Open Data Portal, a platform that gives the general public access to a large amount of the data they collect. We used a data set that provided us the latitude and longitude of all crime occurring in 2015 as well as a description of the crime, date, etc. We used this data along with Chicago's 311 Service Request dataset that collected the location and date of citizen reported vacant lots throughout the city. These two data sets had over 300,00 observations for us to base our analysis on. The final dataset we used was a list of addresses in Chicago from which we sampled locations to compare to the locations with vacant lots.

The city of Chicago was chosen for study because this city's overall crime rate, especially the violent crime rate, is significantly higher than the US average and Chicago contributed for nearly half of 2016's national increase in homicides in the US ("Crime rate in Chicago", n.d.; Sanburn, 2016; Neuhauser, 2016). We hope our study would give some suggestion for both governors and civil engineers in city planning, safety enhancing, and resources allocation.

3. Who will Benefit?

The main group of people that would benefit from this hypothesis is police forces. Being able to more accurately predict crime will allow police forces to set appropriate beats for not just areas of interest but time of day and day of week. This can help with the allocation of resources during a specific time of day and a specific area and not just an area. This can help more directly set a plan to action. We are also looking into which crimes occur more frequently in vacant lots and which happen in non-vacant lots. This will allow police to be better prepared for what they may encounter when they get a call to a vacant lot if a 911 call is unspecific. This could also go into informing police training methods that may help save police officers lives.

Other groups that can be impacted by this analysis are the average people who live in the city. If politicians and public service members can alert people of where to avoid during certain times of days and areas, they can keep their citizens safer for more hours of the day. Also

developers who feel it is not safe to build infrastructure for an underdeveloped community may not flock to these areas. Same goes for businesses in low crime areas. Greenbaum shows that a crime increase in low crime areas drives away businesses leading to more unemployment and thus more poverty and possibly more crime (Greenbaum & Grabmeier, 2005).

4. Objectives and Metrics

Our goal was to try and point out whether or not crimes happen more frequently around vacant lots or non-vacant lots. We accomplished this by creating a 10 meter radius around each vacant building and counting the number of crimes that occurred within that radius in the year of 2015 and comparing that with a similar number of 10 meter radii around non-vacant areas. If we can conclude that there is a higher rate of crime occurring in vacant lots than non-vacant lots at certain times of day then we can provide the police forces with this information.

With this analysis, we were hoping that we would be able to provide local police forces with information that can be used to handle the allocation of police beats throughout the city. Our goal is to increase the number of crimes that are committed in cities and we feel one way to do that is to patrol areas with high crime frequencies more often than areas without a high frequency of crime if resources are limited. Ideally, we would like all areas of the city to be patrolled equally as well at all times, but city budgets have limitations and our goal is to help these funds go as far as possible.

We will be able to measure if our analysis is correct if a couple of things happen. If a police force were to change their beats to match the locations and times of our analysis, we believe that these forces will be able to lower crime in their precincts. Our method can be run again after a few months to determine if crime as a whole has gone down in these areas or times and also overall. Again, our goal is to continue to evaluate methods to help guide the allocations of funds and manpower to try and decrease crime as a whole in the city of Chicago.

5. Methodology

5.1 Kernel Density Estimator

We began by creating Kernel Density Estimators for both vacant lots in the city of Chicago and also the location of crimes. We wanted to see if by these two density plots, we would be able to see a correlation or overlap in the amount of occurrences in similar locations. As one can tell by looking at the KDE's below, there is no overlap in any of the brighter regions. So we knew that KDE would not be an accurate representation of this spatial process.

5.2 Distance Between Crime and Vacant Lots

Next we wanted to investigate whether the distance between crime occurring around a vacant lot was significant or not and if any of these occurrences actually took place. We were able to establish around 3,000 instances of crime happening around vacant lots. With this information, we were able to create 10 meter search radii around each vacant lot and count the

total number of crimes that were committed in that area. In order to have something to compare these areas to, we decided to geocode a random sample of addresses from the list of Chicago addresses and label them as “non-vacant”. We randomly sampled a set of addresses from the list of Chicago addresses in order to fulfill the assumption that our samples were independent. With these vacant and non-vacant address, we would then be able to compare occurrences of crime in both vacant areas and non-vacant areas.

In order to create our set of negative (non-vacant) addresses to compare with our positive (vacant) addresses from our crime observations, we used geocoding code created by Shane Lynn and the Google Maps API. This service provided us with over 6,000 address that were not within a 10 meter radius of crime occurring near a vacant lot. We took a random sample of 3,763 from these negative points in order to match with the 3,763 vacant lot address we calculated above. Figure 1 displays the locations we utilized in our analysis.

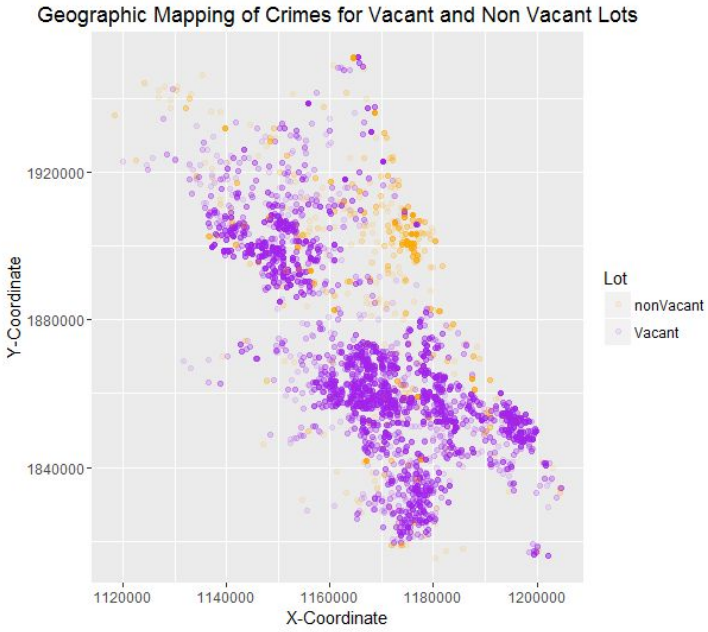

Figure 1

Next we took a count of how many crimes occurred in the areas we determined as vacant lots and the areas of their negatives or non-vacant lots. We found that our sample of random address with non-vacant lots had significantly less crime occurring around them than those lots that were deemed vacant. Since these counts were so different we decided to use Welch’s two sample t-test to test if the difference in mean crimes within a 10 meter radius for vacant lots was different than the mean for non-vacant lots. This leads to our first hypothesis:

5.3 Hypotheses

5.3.1 Hypothesis 1:

- $H_0 : \mu_n = \mu_v$
- $H_A : \mu_n - \mu_v < 0$

Where μ_n refers to the mean number of crimes occurring within 10 meters of each non-vacant address and μ_v refers to the mean number of crimes occurring within 10 meters of each vacant lot.

We then attempted to determine if the distribution of times at which crimes were committed differed significantly between crimes near vacant lots and crimes near non-vacant lots. We created bins for each crime based on day and night, weekend and weekday, and seasons. The bins for day and night were crimes occurring between 1000 and 1600 (day) and crimes occurring between 2200 and 0400 (night). The bins for weekend and weekday were crimes occurring between Monday and Thursday (weekday) and crimes occurring between Friday and Sunday (weekend). The bins for seasons were crimes occurring between December and February (Winter), crimes occurring between March and May (Spring), crimes occurring between June and August (Summer), and crimes occurring between September and November (Fall). We used a χ^2 goodness of fit test for each set of bins. The hypotheses for each test were similar:

5.3.2 Hypothesis 2:

- $H_0 : p_{v1} = p_{n1}, p_{v2} = p_{n2} \ldots, p_{vk} = p_{nk}$
- H_A : *one equation in* H_0 *is not true*

Where k refers to the number of bins in each test and p_{vk} and p_{nk} are the proportion of the kth bin to the total number of crimes in the vacant sample and non-vacant sample respectively.

Lastly, we attempted to determine whether the distribution of types of crime were significantly different in proximity to vacant lots and non-vacant lots. We wanted to test types of crime which were sufficiently represented in the Chicago crime dataset for 2015 so we only tested types of crime which were represented at least 10,000 times in the dataset. These types of crime were, assault, battery, burglary, criminal damage, deceptive practice, motor vehicle theft, narcotics, other offense, robbery, and theft. A χ^2 goodness of fit test was used to test this hypothesis:

5.3.3 Hypothesis 3:

- $H_0 : p_{v1} = p_{n1}, p_{v2} = p_{n2} \ldots, p_{vk} = p_{nk}$
- H_A : *one equation in* H_0 *is not true*

Where k refers to the number of types of crime and p_{vk} and p_{nk} are the proportion of the kth type of crime in the vacant sample and non-vacant sample respectively.

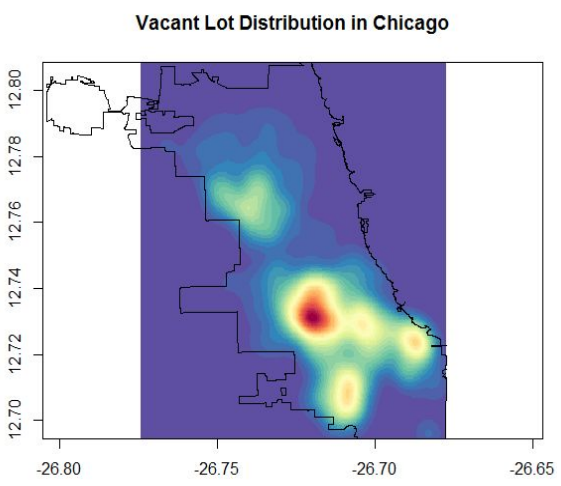

Figure 2

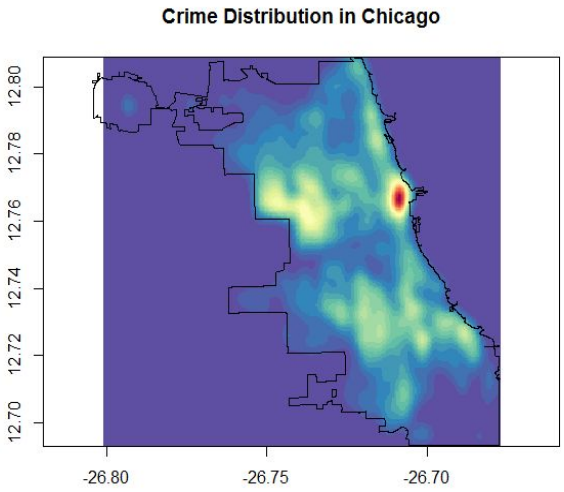

Figure 3

6. Results

6.1 Hypothesis 1: *Rate of crime*

Of the 2015 crimes we examined, 10,195 occurred within 10 meters of our sampled vacant lots. Only 1,219 occurred on our sample of non-vacant lots. $\mu_n = .324$ and $\mu_v = 2.904$. The Welch two sample t-test yielded a close-to-zero p-value of 2.2e-16 when the means of crimes near vacant lots and means of crimes near random address are compared. This means the difference in means of crimes near vacant lots and non-vacant lots is large, thus there is a high likelihood that the rate of crimes near vacant lots in Chicago is higher than the rate of crime near addresses which are not vacant.

6.2.1 Hypothesis 2a: *Time of Day (Day vs. Night)*

An examination of all 2015 crimes revealed that approximately 19.2% occurred at night (2200 - 0400 hours), while 34.9% occurred during the day (1000 - 1600 hours). We picked these hour blocks because no matter what time of year it is, it will always be dark between 2200 and 0400 and it will always be light out during the hours of 1000 and 1600. To test if the distribution of the time of day crimes were committed is different between vacant and non-vacant lots, we conducted a Pearson's Chi-Squared Test ($\alpha = 0.05$). We reject the null hypothesis, seeing that there is a difference in the distribution of crime over time of day between vacant and non-vacant lots ($\chi^2 = 9.8638$, df=1, $p < 0.05$). The percentage of crimes on vacant lots that occurred during the day is significantly higher than the percentage of crimes on non-vacant lots that occurred during the day (see Table 1 below).

Table 1.

	Non-Vacant		Vacant	
	n	Rel. *f* (%)	*n*	Rel. *f* (%)
Day (1000-1600)	387	57.42%	3,546	63.69%
Night (2200-0400)	287	42.58%	2,022	36.31%

6.2.2 Hypothesis 2b: *Day of Week (Weekday vs. Weekend)*

Across all crimes that occurred in 2015, about 56.6% took place on a weekday. We defined weekday as Monday through Thursday, and weekend as Friday through Sunday. To test whether this proportion was consistent across location type, we conducted another Pearson's Chi-Squared test (α=0.05). The Chi-Squared test allows us to reject the null hypothesis that the occurrence of crimes by weekday is independent of location-type (χ^2=6.8713, df=1, p<0.05). As seen in **Table 2.**, the relative proportion of crimes that take place on a weekday is greater for vacant (57.30%) than for non-vacant lots (53.32%).

Table 2.

	Non-Vacant		Vacant	
	n	Rel. *f* (%)	*n*	Rel. *f* (%)
Weekday	650	53.32%	5,842	57.30%
Weekend	569	46.68%	4,353	42.70%

6.2.3 Hypothesis 2c*:* *Season of Year (Spring, Summer, Fall, Winter)*

An examination of all crimes in 2015 revealed that they did not occur equally across the four seasons. Approximately 28% took place in the summer, while 23% too place in the winter. To test whether this proportion was consistent across location type, we conducted another Pearson's Chi-Squared test (α=0.05). We accept the null hypothesis that the occurrence of crimes by weekday is independent of location-type(χ^2=2.136, df=3, p=0.54).

6.3 Hypothesis 3: *Type of Crime*

An examination of all crimes in 2015 revealed that the proportion of types of crimes differed between vacant and non-vacant lots. We only looked at the top 10 most frequently reported crimes since each of these crime types had at least 10,000 reports and the rest of the crime types had significantly less. Deceptive practice and theft constituted a higher proportion of the crimes that occurred on non-vacant lots than of the crimes that occurred on vacant lots (8% and 31% respectively). To test whether this proportion was consistent across location type, we conducted another Pearson's Chi-Squared test (α=0.05). The Chi-Squared test allows us to reject the null hypothesis that the type of crime is independent of location-type (χ^2=297.7, df=9, p<0.05). Table 3, shown below, illustrates the relative proportion of most crimes is higher in vacant lots than non-vacant lots which was expected.

Table 3.

	Non-Vacant		**Vacant**	
	n	Rel. *f* (%)	*n*	Rel. *f* (%)
Assault	80	7.27%	840	9.02%
Battery	213	19.36%	2,380	25.56%
Burglary	52	4.27%	658	7.07%
Criminal Damage	119	10.82%	1,368	14.70%
Deceptive Practice	88	8.00%	341	3.66%
Motor Vehicle Theft	44	4.00%	432	4.64%
Narcotics	53	4.82%	701	7.53%
Other Offense	53	4.82%	832	8.93%
Robbery	53	4.82%	452	4.85%
Theft	345	31.36%	1,309	14.06%

7. Conclusion

These results show that the rate of crime, time of crime, and type of crime are significantly different in proximity to vacant lots as opposed to any other address in Chicago. Given crime's higher frequency near vacant lots, law enforcement can prioritize patrolling areas with a high density of vacant lots if they want to catch more criminals or keep the population

living near vacant lots safer. Since the distribution of crimes in vacant lots was skewed towards the daytime and during weekdays, if forced to make a choice of when to patrol vacant lots, these factors could inform law enforcement to do so during these times. Our results also showed the distribution of the type of crimes was significantly different for vacant lots versus non-vacant lots. Law enforcement could infer that if they were seeking out crimes other than deceptive practice and theft, that vacant lots would be more likely to exhibit those types of crimes proportionally.

8. Limitations and Further Research

We only applied this method to reported crime and reported vacant lots for the year 2015. A further application of this method would be to look across several years or to compare several years to see if trends in this pattern are noticed for more than just one year. Being able to predict these trends and present a police force with this information for several years would provide a more persuasive argument for them to adopt these new data driven conclusions.

A possible flaw in our methodology is we are uncertain whether or not crime happened at a location before or after it was reported vacant. A crime might have occurred in January at an occupied building and that same building may have been reported vacant 6 months later in July. These things we did not take into account due to the size of the data. We feel that by limiting our vacant lot reports and crime to one year significantly cuts down on the possibility of this occurring but it does not take out all these cases.

Furthermore, to classify addresses as non-vacant we made the assumption that our vacant property dataset is comprehensive. We randomly sampled Chicago addresses and removed any that overlapped with our set of vacant properties. If there were vacant properties missing from our dataset, it is possible that our non-vacant addresses overlap with those missing vacant properties.

Data Sources:

311 Service Requests - Vacant and Abandoned Buildings Reported. Retrieved from, https://data.cityofchicago.org/Service-Requests/311-Service-Requests-Vacant-and-Abandoned-Building/7nii-7srd

Crimes - 2015. Retrieved from, https://data.cityofchicago.org/Public-Safety/Crimes-2015/vwwp-7yr9

Chicago Street Names. Retrieved from https://data.cityofchicago.org/Transportation/Chicago-Street-Names/i6bp-fvbx

References:

Blau, J. R., & Blau, P. M. (1982). The Cost of Inequality: Metropolitan Structure and Violent Crime. *American Sociological Review, 47*(1), 114.

Branas, C. C., Rubin, D., & Guo, W. (2012). Vacant Properties and Violence in Neighborhoods. *ISRN Public Health, 2012*, 1-9.

Crime rate in Chicago, Illinois (IL): Murders, rapes, robberies, assaults, burglaries, thefts, auto thefts, arson, law enforcement employees, police officers, crime map. (n.d.). Retrieved September 26, 2016, from http://www.city-data.com/crime/crime-Chicago-Illinois.html

Davis, A. (2016, April 06). In U.S., Concern About Crime Climbs to 15-Year High. *Gallup.* Retrieved August 29 2016, from http://www.gallup.com/poll/190475/americans-concern-crime-climbs-year-high.aspx

Fisher, D. (2015, October 29). America's Most Dangerous Cities: Detroit Can't Shake No. 1 Spot. *Forbes Magazine.* Retrieved August 29, 2016, from http://www.forbes.com/sites/danielfisher/2015/10/29/americas-most-dangerous-cities-detroit-cant-shake-no-1-spot/#29f1501a12c8

Greenbaum, R., & Grabmeier, J. (2005, March 30). Violent Crime Surges Hurt Businesses Most In Low-Crime Areas. *Ohio State Research*. The Ohio State University. Retrieved August 29, 2016, from https://researchnews.osu.edu/archive/crimsurg.htm

Holden, M., & Serpas, R. (2016, August 11). Don't buy the hype: The U.S. is not experiencing a terrible new crime wave. *Washington Post*. Retrieved August 29, 2016, from https://www.washingtonpost.com/opinions/dont-buy-the-hype-the-us-is-not-experiencing-a-terrible-new-crime-wave/2016/08/11/940cda08-5e5e-11e6-9d2f-b1a3564181a1_story.html?utm_term=.eba15202190c

Kraut, D. T. (1999). Hanging out the no vacancy sign: eliminating the blight of vacant buildings from urban areas. *New York University Law Review*, vol. 74, no. 4, p. 1139, 1999.

Neuhauser, A. (2016, September 19). Chicago Driving Uptick in Murders; National Crime Rate Stays Near 'Historic Lows'. *U.S. News & World Report.*

Retrieved September 26, 2016, from http://www.usnews.com/news/articles/2016-09-19/chicago-drives-uptick-in-murders-national-crime-rate-stays-near-historic-lows

Raphael, S., & Winter‐Ebmer, R. (2001). Identifying the Effect of Unemployment on Crime. *The Journal of Law and Economics, 44*(1), 259-283.

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

Sanburn, J. (2016, September 19). Chicago Responsible for Nearly Half of U.S. Homicide Spike. *TIME*. Retrieved September 26, 2016, from http://time.com/4497814/chicago-murder-rate-u-s-crime/

Spelman, W. (1993). Abandoned buildings: Magnets for crime? *Journal of Criminal Justice, 21*(5), 481-495.