Effects of 311 calls for service on the incidence of crime in New York City

Pooja Hiranandani
Computer Science
New York City, US
ph1130@nyu.edu

Abstract—

New York City has experienced a remarkable drop in its crime rate over the past three decades. The credit for this often goes to its adoption of the broken windows theory of policing, though several doubts have been expressed about the efficacy of the theory and its consequences for minorities. I analyze the effects of successfully resolved 311 calls for service related to neighborhood disorder on the incidence of crime at the street segment level, as an evaluation of the soundness of the broken windows theory.

Keywords—analytics, broken windows, 311, crime, new york city

I. Introduction

New York City has experienced a remarkable drop in its crime rate over the past three decades.[1] The credit for this often goes to its adoption of the broken windows theory of policing, though several doubts have been expressed about the efficacy of the theory and its consequences for minorities.[2] The broken windows theory originated from a 1982 *Atlantic Monthly* article of the same name written by James Q. Wilson and George Kelling. They postulated that broken windows and other unchecked social and physical disorder are direct antecedents to criminal behavior. The article led to the development of popular broken windows policing or zero-tolerance policing strategies that targeted minor offenses, such as panhandling, public drunkenness, and vandalism as a mechanism to reduce more serious offenses.[3] There have been several analyses of the efficacy of the theory, but few using big data. To address this gap, I analyze the effects of successfully resolved 311 calls for service related to visible neighborhood disorder on the incidence of crime at the street segment level, as an evaluation of the soundness of the broken windows theory. 311 calls for graffiti removal, street light replacement, poor street condition etc. can, arguably, be considered a valid measure of neighborhood disorder and represent the first line of defense against crime under the broken windows theory. I train two linear regression models, each with counts of 311 data lagged by different durations. The reason for lagged counts is that the impact of resolving neighborhood issues tends to take time to percolate and cause a noticeable effect on crime.[4]

II. Motivation

Prior studies into the efficacy of the broken windows theory treated punitive and policing actions such as misdemeanor arrests as the key feature of the implementation strategy of the broken windows theory. With data now available of 311 calls relevant to the broken windows theory, we can add another dimension to the evaluation of the theory, one that involves agencies other than the NYPD and studies tactics that seek not to punish individuals but improve quality of life in an area. An additional motivation for the study is that prior studies were conducted at the precinct, census block or tract level. The broken windows theory works on the assumption that a visible reduction in disorder will effect a change in the incidence of crime. Analyzing data at the street segment level allows us to look at a small enough portion of geography in which the visual impacts of an increase or reduction in disorder will have a direct bearing on crime levels.[5]

III. Related Work

The broken windows theory has been a subject of tremendous study and debate among social science researchers, city planners and criminologists. Evaluations of the theory have often arrived at contradictory conclusions. (Kelling and Sousa, 2001) studied NYPD precincts and used regression analysis, observational and interview data to demonstrate a statistically significant and substantial negative relationship between broken windows policing and violent crime.[6] (Corman and Mocan, 2005) used monthly time-series data between 1974 and 1999 from

New York City and found that the broken-windows hypothesis has validity in the case of robbery, motor vehicle theft, and grand larceny.[6] (Wheeler Andrew, 2017) used big data to test the broken windows theory of crime by examining the relationship between 311 calls for service and crime at the street segment and intersection level in Washington, D.C. Controlling for a set of micro-level covariates as well as unobserved neighborhood-level effects using negative binomial regression models, it found that detritus and infrastructure-related calls for service have a positive, but small effect on crime. The results suggest that 311 calls for service are a valid indicator of physical disorder where available, and the findings partially confirm the broken windows theory. Given the small effects though, reducing physical disorder is unlikely to result in appreciable declines in crime.[5]

Other studies have found little to no effect of the broken windows theory on the incidence of serious crime. (Harcourt and Ludwig, 2006) used data from an experiment carried out by the U.S. Department of Housing and Urban Development known as Moving to Opportunity (MTO). Officials assigned around 4,600 low-income families from communities largely consisting of public housing and characterized by high rates of crime and social disorder to housing in more reputable and high-status neighborhoods (Orr et al. 2003). Assignment to the program was random, allowing identification of the impact of disorder in the new neighborhood on the relocated individuals' criminal activity, without confounding effects of individuals' background. Harcourt and Ludwig (2006) compared the crime rates of those that moved and those that stayed in their lesser neighborhood, and found that the rate of neighborhood order or disorder has no noticeable effect on criminal behavior. [7] (O'Brien and Sampson, 2015) applied an ecometric methodology to two databases from Boston: 1,000,000+ 911 dispatches; and indicators of physical disorder from 200,000+ requests for non-emergency services. Both distinguish between disorder in public and private spaces. A cross-lag longitudinal analysis was conducted using two full years of data (2011-2012). Their results support a social escalation model where future disorder and crime emerge not from public cues but from private disorder within the community. [8]

Finally, (Braga et al., 2015) identified 30 randomized experimental and quasi-experimental tests of disorder policing. Their meta-analysis suggests that policing disorder strategies are associated with an overall statistically significant, modest crime reduction effect. The strongest program effect sizes were generated by community and problem-solving interventions designed to change social and physical disorder conditions at particular places. Conversely, aggressive order maintenance strategies that target individual disorderly behaviors did not generate significant crime reductions. [9]

The analytic in this paper is aiming to carry out an evaluation of the soundness of the broken windows theory similar to the ones described above but by using big data that can be finessed to mine insights from the smallest units of spatial analysis possible - the street segment.

IV. DATASETS

The first dataset used is the 311 calls for service request available on NYC Open Data. 311 is a telephone number that provides access to non-emergency municipal services. Examples of service requests include abandoned vehicles in roadway, graffiti removal, illegal burning, debris in roadway, code and housing violation etc.[10] The dataset contains 311 Service Requests from 2010 till the present day. It includes details such as the building from which the compliant was made, type of complaint, action taken, date of resolution, etc. [11]

This dataset was ~ 10 GB in size as in July 2018 and was downloaded as a single dump. Data from January 2015 to December 2016 was used for analysis. A total of 869,122 and 836,564 311 calls for service related to visible neighborhood disorder were made and successfully resolved in 2015 and 2016 respectively. The table below presents a non-exhaustive list of the counts of these calls.

311 Complaint Type	2015	2016
Sewer	61428	62604
Street Condition	233868	170110
Sanitation Condition	62080	65810
Sidewalk Condition	53202	52508
Street Light Condition	92268	87872
Overgrown Tree/Branches	21372	18900
Graffiti	24806	23352
Standing Water	2800	11516
Derelict Vehicle	42844	57448
Vacant Lot	3638	3620

The second dataset used is the NYPD Complaints dataset available on NYC Open Data. This dataset includes all valid felony, misdemeanor, and violation crimes reported to the New York City Police Department (NYPD) from 2006 onwards to the most recent completed quarter. It contains information such as date, time, location, type of alleged crime etc. [12] This dataset was ~ 1.5 GB in size as in July 2018 and was downloaded as a single dump. Data from January 2017 to December 2017 was used for analysis. A total of 316,602 misdemeanors and violations were committed in 2017. The table below presents a non-exhaustive list of the counts of different types of misdemeanors and violations committed.

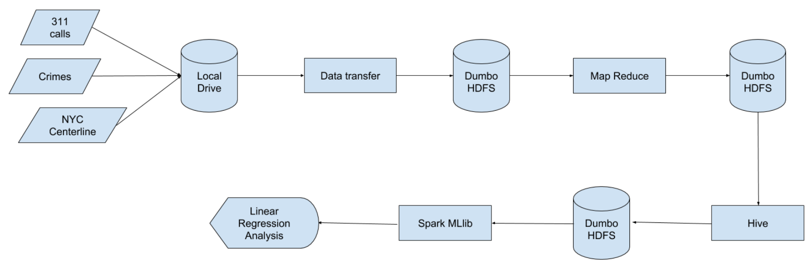
Misdemeanor/Violation Type	**2017**
Petit larceny	82737
Offenses against public order, public sensibilities and the right to privacy	21347
Dangerous Drugs	17181
Assault 3 and related offenses	51020
Criminal trespassing	3590
Prostitution and related offenses	32
Criminal mischief and related offenses	38510
Harassment 2	65590

There were a total of 136,832 felonies committed in New York City in 2017. The table below presents the counts of the seven major felony types.

Felony Type	**2017**
Burglary	11952
Murder and non-negligent manslaughter	268
Grand Larceny	41711
Rape	80
Robbery	13854
Grand larceny of a motor vehicle	5636
Felony assault	19772

The third dataset is the New York City Street Centerline dataset available on NYC Open Data. New York City is broken down into 119,331 street segments with a segment being considered as that portion of the street separated by two intersections, e.g. W 71st Street between Broadway and Westend Ave is one street segment. The dataset was downloaded as a single dump in July 2018 and was ~ 44 MB in size at the time. [13]

V. Design



The project design involved downloading data from the above-mentioned publicly available datasets to my local drive and then transferring the datasets to the HPC Dumbo Hadoop cluster located at Courant Institute of Mathematical Sciences - NYU.[14] The datasets were profiled and cleaned using MapReduce.[15] Then the data was moved to Hive for further refinement and for the mapping of crimes and 311 calls to street segments using the Spatial Framework for Hadoop library.[16,17] Subsequently the data was moved to Spark where two Negative Binomial Regression models were trained and analyzed.[18]

VI. Results

The independent variables in this study are the counts of 311 calls for service related to visible neighborhood disorder at the street segment level in the years 2015 and 2016. The dependent variable is the count of crime committed at the street segment level in the year 2017. Only street segments that had crimes were chosen. Lagged counts of the independent variables were chosen because the effects of resolving neighborhood disorder can take time to have an effect on the incidence of crime in that area. [4] Crimes and 311 calls were mapped to street segments using the Spatial Data Analysis for Hive library. [17] The mapping was initially attempted through a user-defined function which was messy and inaccurate. The discovery of the external library with out-of-the-box functions for distance calculations made the mapping much easier and more accurate.

Negative Binomial Regression was the method of analysis chosen for this study since this type of regression is recommended when the dependent variable is a count variable and the variance of the variable is greater than its mean, which is the case here.[19] Since Spark does not have an out-of-the-box Negative Binomial Regression function, the Generalized Linear Regression function was used with the Tweedie distribution family.[20] A variance power of 2 was set and the log link function was used.[19]

Two models were built - one that involved 2016 counts of the independent variable and the other that involved 2015 and 2016 counts of the independent variables. The results are as follows:

Model 1

Factor	Coefficient	Standard Error	P value
311 calls (2016)	0.02404	0.0007	< .01

Model 2

Factor	Coefficient	Standard Error	P value
311 calls (2016)	0.0153	0.0009	< .01
311 calls (2015)	0.0113	0.0009	< .01

From the results of the regression analysis, it is clear that successfully resolved 311 calls for service related to visible neighborhood disorder at various durations of lag have a positive and statistically significant effect on crime committed on that street segment, however the effects are very modest. Adding more lagged counts of 311 calls did not explain the incidence of crime much more than a single lagged count as the sum of the coefficients is almost the same in both models. This suggests that crime has many more significant determinants than the broken windows theory takes into account. Furthermore, if resolving neighborhood disorder did have a *deterrence* effect on crime as the theory suggests, the coefficients would be *negative,* implying that a increase in the resolution of neighborhood disorder in previous years would be followed by a decrease in the incidence of crime and vice versa.

There could be a few reasons for the results above. Among them could be that while neighborhood disorder is accompanied by high levels of crime, simply resolving neighborhood disorder does not do much to decrease the incidence of crime and both disorder and crime have other causes and require other interventions to be addressed.

Another reason could be that 311 calls for service are a community based form of disorder policing which will, by its nature, be reactive and hence disorder that is resolved successfully will, in some cases like overflowing garbage, standing water, be undone relatively quickly without proactive maintenance by the city and will have a meagre effect on crime for that reason. However, resolution of infrastructural related disorder like street conditions, malfunctioning street lights, etc. are more long-lasting and ought to have an impact on crime in the following months and years, as per the theory.

It is possible that there are other pertinent variables missing from the analysis that need to be controlled for such as socio-economic characteristics of the street segment, locational characteristics of the street segment like whether it is near a subway station or a bar or whether it is a segment that has a lot of foot traffic or road traffic etc. and misdemeanor arrests which is the second line of defense against serious crime as per the theory.

It is also possible that while the broken windows theory of crime reduction is not correlated to a decrease in crime in the current day, it did have an impact on the crime levels in New York City in the 1990s and the reductions achieved then represent the limits of the theory.

VII. Future Work

There is scope for further research into the broken windows theory using big data sources such as the ones I used in this analytic. It would be interesting to see if lagged counts of 311 calls for service have an impact on present counts of 311 calls since this effect is also implied by the broken windows theory - that a reduction in disorder breeds order. Additionally I might have liked to include some control variables in the regression equation like the locational characteristics of the place such as whether it is near a highway or subway station, whether there are bars or shopping centers nearby etc. Additional controls like the number of police-persons and parking enforcement officers assigned to patrol the area under consideration and the socio-economic profile of the area would have made sense to include. Another interesting area of research would be the effects of parking enforcement itself on serious crime - whether the physical presence of parking enforcement officers and the enforcement of a minor violation like a parking offense has a statistically significant effect on more serious crime.

VIII. Conclusion

New York City has been center stage in policy and scholarly debates about policing disorder and the broken windows perspective (most recently, see Rosenfeld, Terry, and Chauhan 2014; Zimring 2012). Although local officials and national observers attribute the city's violent crime drop in the 1990s to the adoption of the broken windows policing strategy, many academics argue that it is difficult to credit this specific strategy with the surprising reduction in violent crime. The New York Police Department (NYPD) implemented the broken windows strategy within a larger set of organizational changes framed by the Compstat management accountability structure for allocating police resources (Silverman 1999). As such, it is difficult to disentangle the independent effects of disorder policing relative to other strategies implemented as part of the Compstat process (Weisburd et al. 2003). Other scholars suggest that a number of rival causal factors, such as the decline in the city's crack epidemic, played a more important role in the crime drop (Bowling 1999). Some academics have argued that the crime rate was already declining in the city before the implementation of police reforms and that the city's decline in homicide rates was not significantly different from declines experienced in surrounding states and in other large cities that did not implement aggressive enforcement policies during that time period (Baumer and Wolff 2014; Eck and Maguire 2006). [9]

My analytic using big data confirms much of the above mentioned skepticism regarding the efficacy of the broken windows theory. It suggests that there are many more significant determinants of crime than public disorder. A recent analysis using big data sources similar to mine, conducted in Washington D.C., similarly concludes that neighborhood disorder has little impact on the incidence of crime. [5]

While many studies, including mine, have expressed doubts about the soundness of the broken windows theory of policing, the theory still has a lot of defenders, perhaps at the cost of introducing more effective means of reducing crime. It is therefore important that we continue to keep the research spotlight on this theory, to guide and influence the very significant ways in which our cities and countries are designed and manipulated in the quest to address and reduce crime.

Acknowledgments

The HPC team at Courant Institute of Mathematical Sciences, NYU and Prof. Suzanne McIntosh

The creators of the Spatial Framework for Hadoop Library https://github.com/Esri/spatial-framework-for-hadoop/tree/master/hive/src/main/java/com/esri/hadoop/hive

NYC Open data https://opendata.cityofnewyork.us/data/

Google Cloud Platform https://cloud.google.com/

References

1. Langan; Matthew R. Durose (December 3–5, 2003). "The Remarkable Drop in Crime in New York City". 2003 International Conference on Crime. Rome, Italy.
2. T. White. Hadoop: The Definitive Guide. O'Reilly Media Inc., Sebastopol, CA, May 2012.
3. Maskaly, J., & Boggess, L. N. (2014). Broken Windows Theory. The Encyclopedia of Theoretical Criminology, 1-4.
4. Corman, H., & Mocan, N. (2002). Carrots, Sticks and Broken Windows.
5. Wheeler, A. P. (2017). The Effect of 311 Calls for Service on Crime in D.C. at Microplaces. Crime & Delinquency, 001112871771497.
6. L. Kelling, George & Sousa, William. (2001). Do Police Matter? An Analysis of the Impact of New York City's Police Reforms.
7. Weele, J. J., Flynn, M. P., & Wolk, R. J. (2017). Broken Window Effect. Encyclopedia of Law and Economics, 1-4.
8. O'Brien, D. T., & Sampson, R. J. (2015). Public and Private Spheres of Neighborhood Disorder. Journal of Research in Crime and Delinquency, 52(4), 486-510.
9. Braga, A. A., Welsh, B. C., & Schnell, C. (2015). Can Policing Disorder Reduce Crime? A Systematic Review and Meta-analysis. Journal of Research in Crime and Delinquency, 52(4), 567-588.
10. https://en.wikipedia.org/wiki/3-1-1
11. https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9
12. https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i, https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Current-YTD/5uac-w243
13. https://data.cityofnewyork.us/City-Government/NYC-Street-Centerline-CSCL-/exjm-f27b

14. S. Ghemawat, H. Gobioff, S. T. Leung. The Google File System. In Proceedings of the nineteenth ACM Symposium on Operating Systems Principles – SOSP '03, 2003.
15. J. Dean and S. Ghemawat. MapReduce: Simplified data processing on large clusters. In proceedings of 6th Symposium on Operating Systems Design and Implementation, 2004.
16. Thusoo, Ashish & Sen Sarma, Joydeep & Jain, Namit & Shao, Zheng & Chakka, Prasad & Zhang, Ning & Anthony, Suresh & Liu, Hao & Murthy, Raghotham. (2010). Hive - A Petabyte Scale Data Warehouse Using Hadoop. Proceedings - International Conference on Data Engineering. 996-1005. 10.1109/ICDE.2010.5447738.
17. https://github.com/Esri/spatial-framework-for-hadoop/tree/master/hive/src/main/java/com/esri/hadoop/hive
18. Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S. and Stoica, I. (2010) Spark: Cluster Computing with Working Sets. Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing (HotCloud'10), USENIX Association, Berkeley, 10 p.
19. Kokonendji, Célestin & Demétrio, Clarice & Dossou Gbete, Simplice. (2004). Overdispersion and Poisson-Tweedie exponential dispersion models. VIII Journées Zaragoza-Pau de Mathématiques Appliquées et de Statistiques :Jaca, Spain, September 15-17, 2003, 2003-01-01, ISBN 84-7733-720-9, pags. 365-374. 31.
20. https://spark.apache.org/docs/2.2.0/api/java/org/apache/spark/ml/regression/GeneralizedLinearRegression.html