How Well Does Crime Predict Crime?

A Time Series Analysis of Recent Chicago Crime Data

by

Benjamin Milner

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Art

Queen's University

Kingston, Ontario, Canada

August 2013

Acknowledgments

I would like to express my sincere thanks to Charles Beach. His knowledge and insights have been truly invaluable throughout this process, and his patience and good humor have been much appreciated. As well, I would like to thank my Queen's Economics classmates, many of whom sacrificed considerable time and energy both assisting and entertaining me during my Master's education. I wish to acknowledge my parents, who have always supported and when needed gently pushed me to achieve my goals. Finally, I wish to acknowledge my Lord Jesus Christ for giving me the strength and ability to reach this point.

Table of Contents

List of Tables and Figures

Introduction

It has long been obvious that people desire and perhaps need safety. In Cicero's *On the Laws,* he stated that "the safety of the people is the supreme law." When Abraham Maslow constructed his famed "hierarchy of needs" in 1943 – in which he ranked human needs from most basic to least – he placed safety (defined as security of body, resources, possessions, and family) second only to physiological needs such as breathing and eating. More recently, opinion polls have consistently pointed to safety, in particular safety from crime, as a chief concern of the general population. For example, a March 2013 Gallup poll of Americans showed that 48% worried "a great deal" about crime and violence in America, and a further 27% worried "a fair amount".[1]

Thus, given its importance, one can understand the population's general anxiety when safety from crime seemed to be slipping away in America in the second half of the 20th century: the property crime rate nearly tripled between 1960 and 1991, while the violent crime rate more than quadrupled.[2] In order to counteract this rise, police forces developed several new crime fighting techniques. Chief among these was an increased use of crime statistics. In particular, the New York City Police Department developed a management strategy known as CompStat, in which crime statistics (beginning simply with pinned crime locations on a map, but

[1] The question was phrased as follows: *"Next, I'm going to read a list of problems facing the country. For each one, please tell me if you personally worry about this problem a great deal, a fair amount, only a little, or not at all. First, How much do you personally worry about crime and violence?"*

[2] According to the FBI's Uniform Crime Reports as prepared by the National Archive of Criminal Justice Data, the national number of property crime incidents per 100,000 people in 1960 was 1,726.3, and the equivalent violent crime statistic is 160.9. In 1991, the rates were 5,140.2 and 758.2, respectively.

gradually becoming more sophisticated[3]) played a major role in identifying problem locations and potential trends. Today, many believe that CompStat was pivotal to New York's huge drop in crime in the 1990s.[4] As such, it has been widely copied, with a 2004 survey reporting that 58% of large (100 or more officers) American police forces were using or planning to use in the near future some form of CompStat.[5] Meanwhile, crime across the United States has fallen substantially from its peak in the early nineties: the violent crime rate fell from 758.2 incidents per 100,000 people in 1991 to only 386.3 in 2011, its lowest since 1970, while the property crime rate fell from 5140.2 in 1991 to 2908.7 in 2011, its lowest since 1967. It is worth noting that Canadian crime rates have followed similar trends, peaking in the early nineties and falling substantially since.[6]

Thus, a common philosophy in modern policing is prevention through prediction. Across the globe, crime-prediction software is changing the way police forces conduct themselves, as they recognize that clear benefits in crime reduction appear to be had from better understanding the dynamic patterns of crime. With that in mind, this project seeks to further study how the frequencies of specific types of crime are influenced by time, each other, and income. The crimes studied in this paper are burglary, robbery, motor vehicle theft, arson, and battery. These were specifically chosen mainly so as to lessen the impact of reporting bias in

---

[3] Douglas Werner Perez and Michael Barkhurst, *Paradoxes of Leadership in Police Management*, pg. 209.
[4] Vincent E. Henry, *The COMPSTAT Paradigm: Management and Accountability in Policing, Business and the Public Sector*, pg. 4.
[5] Weisburd et al., "The Growth of CompStat in American Policing," Police Foundation (2004).
[6] Comparisons between Canada and America's Uniform Crime Statistics are difficult, given that each country defines different types of crimes in different ways, meaning that what is considered a violent crime in Canada may not be recorded as such in America, and vice versa. That being said, according to the Uniform Crime Reporting Survey published by Statistics Canada, Canada's violent crime rate increased by nearly 400% between 1962 and 1992, from 221 to 1084, and has since fallen significantly, dipping to 866 in 2011. Similarly, the property crime rate rose from 1891 in 1962 to 6160 in 1991, but by 2011 was back down to 2547.

data. The first part of the project will consist of a univariate time series analysis of each of the crimes, so as to get a basic understanding of how each develops in time. The second part will then use multivariate time series analysis to determine how the types of crime influence each other. Forecasts will also be developed for both the univariate and multivariate models and compared so as to determine if using lagged values of other types of crime helps in prediction. Finally, I will attempt to determine the impact of income levels on each type of crime by using panel data focused on the community level, again with the hope of improving the ability to predict, and thus prevent, crime.

Literature Review

What follows is a brief description of the academic literature on crime, focusing mainly but not exclusively on the economic perspective, along with discussion on how it relates to this paper. Surprisingly, given the importance of safety from crime, the economics of crime is a relatively new area of research, coming into being only in the last half century. It is generally agreed that the field was birthed by Gary Becker with his 1968 paper, "Crime and Punishment: an Economic Approach," in which he attempted to model criminal behaviour by merely focusing on the incentives and disincentives associated with crime. According to his model, four main factors must be considered: the possible utility gain of a crime, the severity of the punishment if caught, the chance of getting caught, and the opportunity cost of not participating in other work. As stated above, Becker's analysis spawned a whole field's worth of economic inquiries concerning crime, and much of this subsequent work focused on one or more of these factors.

For example, concerning the impact of punishment severity, there exists a sizable literature weighing the impact of the death penalty in terms of deterring crime.[7] Much has also been published about how “three-strike” legislation affects crime rates.[8] More recently, social capital has been recognized within economics as an important factor of crime: the social stigma attached to crime can be viewed as a cost, and thus the degree to which an individual’s peers accept/stigmatize crime becomes relevant. The importance of social factors is made clear through papers such as Freeman (1986), which found that church attendance was strongly correlated with lower crime levels among disadvantaged youth, and Case and Katz (1991), which examined the relation between an individual’s likelihood to commit crime and the prevalence of crime among their peers.

This concept of social capital is implicitly a key feature of the “Broken Window Theory”, introduced by Kelling and Wilson in a 1982 *Atlantic Monthly* article and now a popular concept in contemporary criminology.[9] Loosely speaking, the theory has it that a known prevalence of petty crimes, such as vandalism, can change the social norms within a community concerning crime, making more serious crimes appear more acceptable. This suggests that the frequency of crime in a community will show persistence over time. Indeed, Fajnzylber et al. (1998), an

---

[7] The classic work on this subject is Ehrlich (1975), in which he used empirical analysis to argue that the death penalty did significantly deter crime. Since then much has been published on the matter, along with a great deal of evidence both supporting and refuting Ehrlich’s original conclusion. Yang and Lester (2008) offers a good survey of this literature.

[8] A good recent example is Helland and Tabarrok (2007).

[9] “Broken Windows,” *Atlantic Monthly*, March 1982.

empirical study of the determinants of crime using Latin American data, found that changes in national crime rates tended to persist long after the initial causes of change had vanished, and labeled this effect "criminal inertia." Thus, there is both theoretical and empirical support for this paper's use of lagged values to help forecast crime, and indeed, the broken window theory suggests that lagged values of various types of crime could be useful in forecasting each other, providing a basis for the forthcoming multivariate analysis.

Concerning the opportunity cost of crime, Becker's model is clear that an improvement in legal opportunities should negatively impact crime. One would thus expect higher wages to be correlated with lower crime rates. And indeed, numerous studies at the individual level have found just this effect. Notably, Myers (1983) demonstrated that the likelihood of released property criminals reoffending is strongly deterred by higher legal wages. However, this has been difficult to demonstrate empirically at the community level. One possible reason for this is that a community with a higher average wage may also offer higher returns to crime.[10] The higher gains from crime associated with greater community wealth may increase the crime rate at the same time as the higher wage rate is decreasing it, making the overall effect difficult to predict. Similarly, some empirical studies have argued that, counter intuitively, an increase in a region's employment rate is positively related to an increase in crime.[11] Cantor and Land (1985) hypothesised that the reason for this is that higher unemployment leads to fewer opportunities to commit property crime, particularly because more people stayed at home. Taken together,

[10] Doyle, Ahmed, Horn (1999), 719.
[11] Notably, Cantor and Land (1985), Trumbull (1989).

the literature's ambiguity means that it is difficult to make an a priori hypothesis concerning this paper's analysis of crime using income variables.

One may have already noted that much of the discussion thus far has concerned property crime. While it is perhaps true that property crime lends itself more readily to economic analysis, it is worthwhile to spend some time discussing the factors relating to violent crime, particularly given that the forthcoming analysis includes several violent crimes among its studied variables. One may assume that economic variables such as income and employment would have little to do with violent crimes, which are often viewed by society as motivated by passion and emotion. However, Fajnzylber et al.'s 2002 cross-country study of the causes of homicides and robberies demonstrated that economic variables such as income inequality and GDP growth are statistically significant predictors of both homicides and robberies. It also demonstrated that both these crimes showed noted persistence, providing further justification for this paper's use of time series data. That being said, it seems likely that the violent crimes studied by this paper will be less correlated with economic variables than the property crimes considered.

Data Used in Study

The primary dataset used in this study is the City of Chicago Data Portal series on crime. Similar to the national pattern discussed earlier, Chicago crime has fallen steadily since the early nineties. For example, like the national rate, Chicago's property crime rate peaked in

1991 at 8227.6 incidents per 100,000 people, but had fallen to 4373.2 by 2011.[12] Given that Chicago crime rates have broadly followed the national patterns, extrapolating the results beyond the city scope does not seem inappropriate. First made available to the public in late 2011, this dataset contains detailed information concerning the description, date (down to the hour), and location (down to the block level) of every crime reported in Chicago from the beginning of 2001 (with the exception of murder cases deemed sensitive). It is current up to approximately one week ago, and is updated daily from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. It should be noted that this dataset only contains crimes within the city of Chicago proper, which in 2011 had an estimated population of approximately 2.7 million.[13] It should also be noted that some data points are occasionally modified in light of additional information; in particular, the type of crime described can change as more evidence becomes available. Thus, in order to lessen the possibly distortionary effects of misreported crimes, this paper makes use data only up to the end of 2012, as these crimes have most likely been thoroughly investigated, making misreporting their type less likely. Another possible concern is that some crimes may go unreported, and thus the dataset would not be an accurate reflection of the true crime rate. In particular, low value property crime and crimes such as rape which can impart social stigma on the victim often go unreported.[14] As well, one would expect so-called “crimes against society”,

---

[12] As described in the FBI’s Uniform Crime Reports. It is difficult to compare Chicago’s violent crime rate with the national rate, as the data collection methodology used by the Chicago police force to record instances of forcible rape does not meet UCR guidelines. However, again according to the FBI’s Uniform Crime Reports, the homicide and manslaughter rate in Chicago peaked in 1992 at 33.1 and fell to 15.9 in 2011, indicating that Chicago’s violent crime trends also mirrored national ones.

[13] According to the American Community Survey five-year estimate, the population was 2,700,741 with a margin of error of +/- 65 people.

[14] P. Fajnzylber et al (2002), pg. 1326

such as gambling and prostitution, to be less frequently reported, as there is often not a clear victim. In order to minimize these distortions, the crimes chosen for this analysis are those that are likely to be reported given that they are relatively serious, do not in general stigmatize the victim, and they have a clear victim. These crimes are burglary, robbery, motor vehicle theft, arson, and battery. In addition to minimizing reporting bias, this choice of crimes allows one to study both property (burglary, motor vehicle theft) and violent (arson, battery) crime, as well as the intersection between the two (robbery, for example, is essentially violent theft). Given the large variety of forms crime can take, it is important to explicitly define what each type of crime actually refers to. Table 1 provides the Chicago Police Department's definitions of each type of crime studied.

For the purpose of this paper, the total number of reported incidents of each type of crime was aggregated at the monthly level, so as to enable medium term time series analysis. Monthly summary statistics for each crime are shown in Table 2. Several things can be gleaned from this table. First off, arson appears to be the odd man out, given the small amount of observations. Another notable feature of Table 2 is that the monthly standard deviations are not insignificant. This is likely related to the considerable seasonality exhibited by all five crimes. The downward trend followed by most of the crimes (the notable exception is burglary) also plays a role. Both the seasonality and trend of the crimes can be seen in the graphs of Figure 1.

Table 1: Crime Definitions

Crime	Definition
Burglary	"The unlawful entry into a building or other structure with the intent to commit a felony or a theft."
Robbery	"The taking or attempting to take anything of value under confrontational circumstances from the control, custody, or care of another person by force or threat of force or violence and/or by putting the victim in fear of immediate harm."
Vehicle Theft	"The theft of a motor vehicle."
Battery	"A person commits battery if he intentionally or knowingly without legal justification and by any means, (1) causes bodily harm to an individual or (2) makes physical contact of an insulting or provoking nature with an individual."
Arson	"To unlawfully and intentionally damage or attempt to damage any real or personal property by fire or incendiary device."

Source: CLEAR, obtained from http://gis.chicagopolice.org/clearmap_crime_sums/crime_types.html#N05

Table 2: Monthly Summary Statistics by Crime Type

Type	Total No. of Reported Incidents, 2001-2012	Monthly Mean	Monthly Std. Dev.	Monthly Min, Max
Burglary	304,713	2116.063	297.9982	1339, 2707
Robbery	192,039	1333.604	223.9651	693, 1941
Vehicle Theft	249,916	1735.528	327.6419	1074, 2894
Battery	940,155	6528.854	1300.991	3832, 9802
Arson	8643	60.02083	19.28783	22, 125

Source: Author's calculations from CLEAR database.

Figure 1: Monthly Frequency of Chicago Crime, 2001-2012

Source: Author's calculations from CLEAR database.

Also observable in Figure 1 are notable jumps in the frequencies of robberies and vehicle thefts in 2008, near the start of the recession. This seems to suggest that income and employment play important roles in at least these two types of crime. Burglary also appears to display a slight upward trend from 2007 to 2010. However, it is difficult to find any reaction in the arson or battery rates with regard to the recession. This is in line with our previous assumption that economic variables will not have as much of an effect on violent crime frequencies. Of course, these relationships will receive a more detailed analysis in later sections of this study.

As stated previously, the first two empirical sections of the paper conduct analysis at the city-wide level. The time period considered is 2001-2012, and thus contains 144 monthly time periods.

For the final section of the analysis of this paper, the crime data are broken down to the community level and the effect of income is examined. The city of Chicago consists of 77 community areas (a map of these is provided in Appendix I). Community level median household income data were gathered from three surveys: the 2000 United States Census, the 2006-2010 American Community Survey (ACS) five-year estimate, and the 2007-2011 ACS five-year estimate.[15, 16] Given the lack of monthly income data, it was judged best to examine crime frequency at the annual level in this final section. Still, some additional work was required to

[15] It should be noted that each community area spans several census tracts. Thus, the income data used in this paper was collected from analysis done by consultancy Rob Paral and Associates which aggregated the relevant census tracts so as to determine community area median income.

[16] The 2000 Census was the last to gather household economic data. Now the Census Bureau relies on the ACS, which generates one-, three-, and five-year estimates of various economic variables. For low population areas, such as Chicago's community areas, only five-year estimates are calculated. *Using the American Community Survey: Benefits and Challenges* (2007) provides further information concerning the ACS and its methods.

obtain annual median household income estimates for each community. This is described in detail in Appendix II. Unfortunately, consistent location information for each crime is available only for crimes reported after March 2002. Thus, data from 2001-2002 are dropped in this portion of the analysis, meaning that this section analyses 10 annual time periods (2003-2012). However, given that 77 communities are considered, there are a total of 770 annual observations for each crime type. It should also be noted that there are very few crimes within this new time period that do not have their location reported. Given their extreme infrequency[17], and that they appear to be relatively evenly spread across time, it is judged best to simply drop these crimes from the analysis.

Univariate Analysis of Crimes

As previously stated, the first goal of this paper is to identify how each crime changes over time using univariate time series models. However, it is important first to address the seasonality and trend components shown by each crime. Thus, each crime is regressed using eleven monthly dummies (one for every month except December, so as to avoid collinearity) and a linear trend variable. Note that this process was conducted in a single regression, however this is essentially equivalent to running two separate regressions, one for deseasonalizing and a second for detrending, given the Frisch–Waugh–Lovell theorem.[18] When

[17] These crimes represent 0.05% of burglaries, 0.03% of batteries, 0.02% of robberies, and 0.01% of vehicle thefts in the 2003-2012 sample considered.
[18] A good description of this theorem is provided in Davidson and MacKinnon's *Econometric Theory and Methods* (2004), pg. 62-75.

this regression is carried out, the estimated monthly trend is judged to be significant at the 99% level of confidence for all variables except burglary.

Of course, the strength and causes of the trends exhibited by each of the crimes are interesting points to study in and of themselves. While these are not the focus of this paper, a brief description of the trend of each crime is provided here. As previously described, crime rates have been falling across the United States since the mid-nineties for reasons as varied as a drop in the use of crack cocaine to higher rates of abortion.[19] What is perhaps more interesting is that, at least in Chicago, some crimes have dropped off much more quickly than others. Table 3 displays two sets of figures: the "estimated monthly trend" column gives the trend coefficient estimate from the above regression, which can be interpreted as the monthly fall in incidents due to the trend; "change in annual total, 2001-2012" simply indicates the percent change in incidents reported in 2012 compared to 2001, giving a sense of how large an impact the trend had over the period.[20]

Table 3: Strength of Crime Trends

Crime	Estimated Monthly Trend	Change in Annual Total, 2001-2012[21]
Burglary	-0.1151661	-12.48%
Robbery	-2.805774***	-26.93%
Vehicle Theft	-6.088869***	-40.36%
Arson	-0.3593871***	-54.35%
Battery	-22.55466***	-35.41%

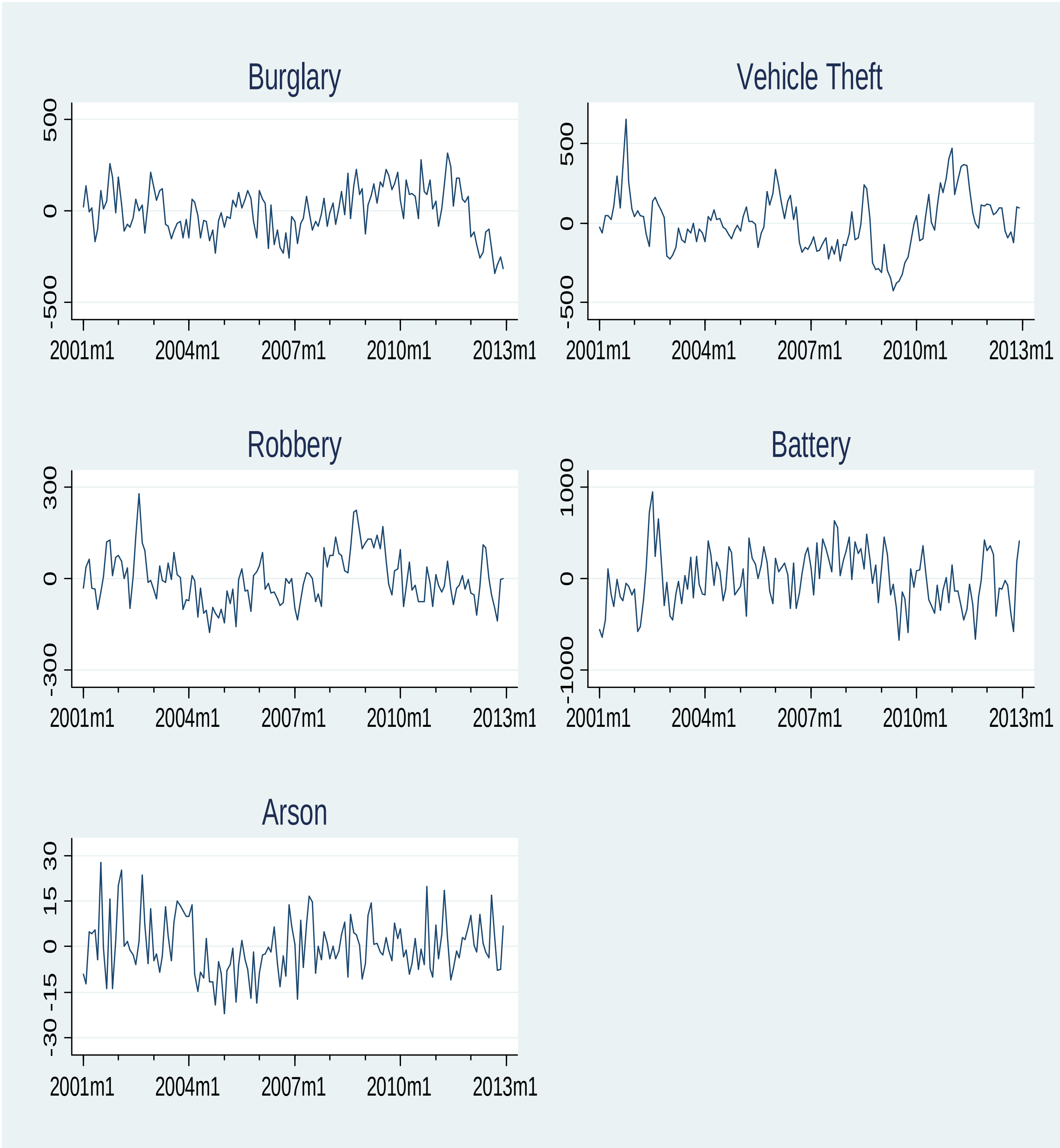
* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database.

[19] Levitt (2004).
[20] The major exception to this is burglary, for which, as can be seen in Figure 2, the crime rate in 2012 has been much lower than predicted by the trend. Thus, -12.48% is much larger than the change predicted by the trend.
[21] {1 – [(total incidents reported in 2012) /(total incidents reported in 2001)] } x 100%.

As the table shows, while all five crimes showed a negative trend, the strength varied greatly across crimes. Arson and vehicle theft have fallen by huge amounts, while burglary has shown a far more modest decrease. One possible explanation for the small drop in burglaries could be that property crimes show more long-term persistence than violent crimes such as robbery, arson, and battery. Meanwhile, the drop in vehicle theft can perhaps partially be explained by advances in vehicle anti-theft technology.[22] This hypothesis may be worth further study in another paper. However, it is the residuals of the above regressions that this section of the paper concerns itself with. Given that burglary does not show a statistically significant trend, the residuals studied for it are only deseasonalized. The remaining crimes are studied using the detrended and deseasonalized residuals, and are shown, along with burglary, in Figure 2.

[22] Farrell et al. (2011) provides a recent analysis of this. A good economic analysis (though slightly dated) is Ayres and Levitt (1998).

Figure 2: Deseasonalized, Detrended Monthly Crime Residuals[23]



Source: Author's calculations from CLEAR database.

[23] With the exception of burglary, which is merely deseasonalized given that there is judged to be no trend to begin with.

Following Box and Jenkins'[24] method, the first step before identifying models is to test each set of residuals for stationarity. To do this the augmented Dickey-Fuller (ADF) test is used.[25] First, the optimal amount of lags to be used for the ADF test is determined by first-differencing each crime, then individually regressing these first-difference variables using different AR(p) processes, working down from AR(13) to AR(0).[26] As suggested by Elliot et al. (1996), the value of p chosen for each variable is that which minimizes the Bayesian Information Criterion (BIC).[27] Then, using these optimal lags, the ADF test is applied to each set of residuals. The results are given in Table 4.

Table 4: Unit Root Test Results

Crime	Optimal Lag Length Used	ADF Test Statistic	Approx. p-value
Burglary	2	-2.847	0.0542
Robbery	1	-4.718	0.0001
Vehicle Theft	1	-3.868	0.0023
Arson	1	-7.539	0.0000
Battery	1	-6.058	0.0000

Source: Author's calculations from CLEAR database.

With the exception of burglary, the null hypothesis of the presence of a unit root is strongly rejected at the one percent level of significance in favour of the alternative hypothesis

[24] George E. Box and Gwilym M. Jenkins, *Time Series Analysis: Forecasting and Control* (1970).
[25] Again, an excellent description is provided in Davidson and MacKinnon (2004), pg. 620-623.
[26] The max lag of 13 was chosen following the commonly observed rule suggested by Schwert (1989) of letting $p_{max} = 12(T/100)^{.25}$.
[27] The Bayesian Information Criterion, developed by Schwarz (1978), is a model selection criterion defined as BIC = $(-2) \cdot \ln(L) + k \cdot \ln(T)$, where L is the max value of the likelihood function, k is the number of parameters being estimated, and T is the number of observations.

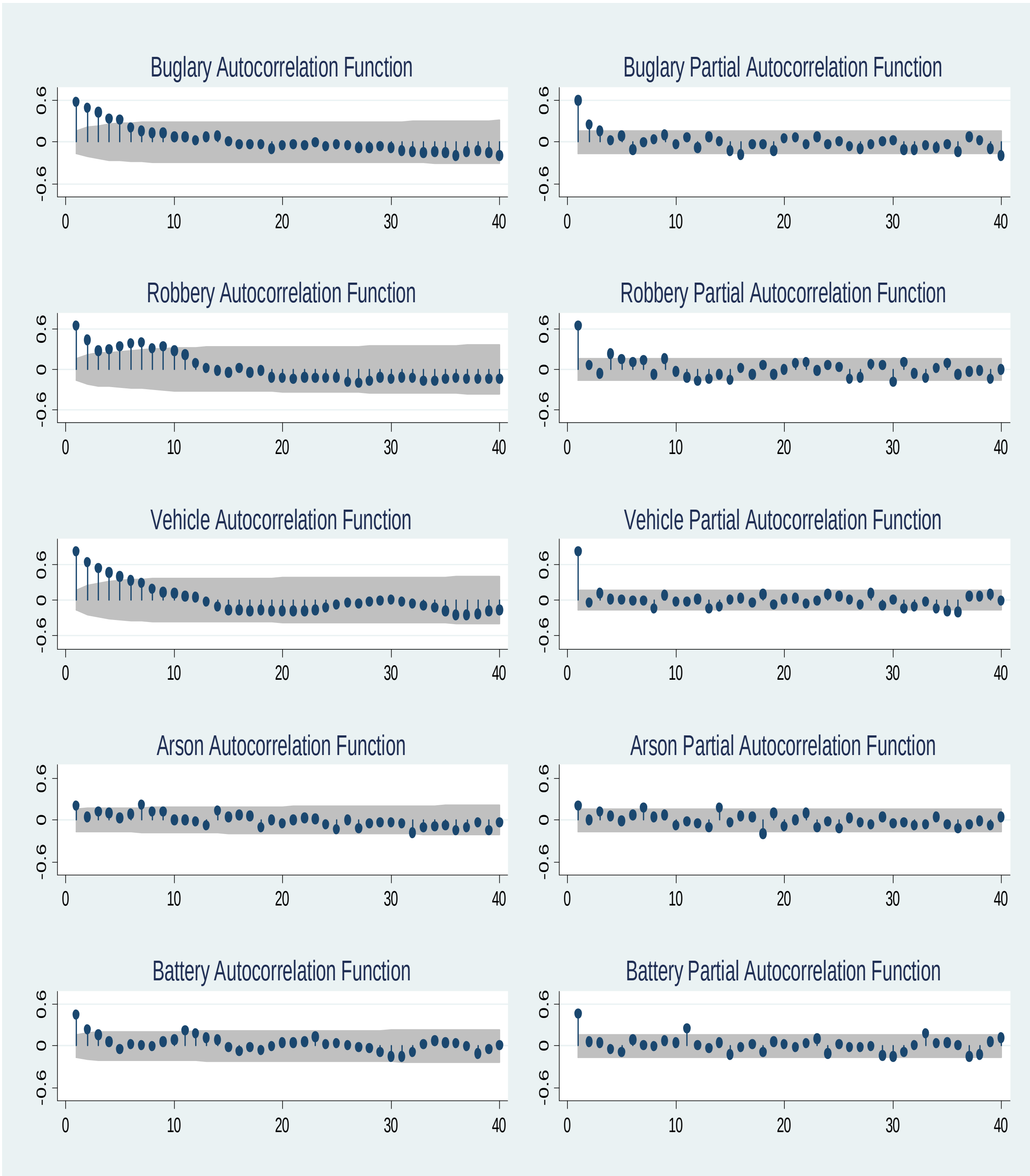
of stationarity for all crimes. For burglary, we can nearly reject the null at the 5% level, and can very safely reject it at the 10% level. This is judged to be sufficient to accept stationarity.

Having shown that each variable is stationary in levels, we proceed to the identification stage. Again following Box and Jenkins (1970), the autocorrelation (AC) and partial autocorrelation (PAC) functions of each variable (shown in Figure 3) are first examined in order to get a rough idea of the short-term persistence demonstrated by each. Using these functions, a max autoregressive lag of p and a max moving average lag of q are chosen for each variable. ARMA models are then constructed and tested down from ARMA(p,q) to ARMA(0,0). The models that minimize the BIC for each crime are then examined further.

Considering burglary first, the first five lags of its AC are judged to be significant at the 95% level of confidence, while the remaining lags up to forty are judged to be insignificant. This suggests that its data generating process contains no more than five moving average lags. Only the first three lags of its PAC are significant at the 95% confidence level, indicating that the DGP contains no more than three autoregressive lags. Thus, every possibility between ARMA(3,5) and ARMA(0,0) is examined. Of the models tested, the ARMA(1,1) yielded the lowest BIC. This chosen model is supported by the fact that we cannot reject, based on the Q-test, that its residuals are white noise, indicating that adding additional lags would not be useful.[28]

[28] The Q-test, also known as the portmanteau white-noise test or the Ljung–Box test, was first described in Ljung and Box (1978).

Figure 3: Autocorrelation and Partial Autocorrelation Functions



Note: For Autocorrelation Functions, Bartlett's formula for MA(q) 95% confidence bands; for Partial Autocorrelation functions, 95% confidence bands [se = 1/sqrt(n)].
Source: Author's calculations from CLEAR database.

This same procedure was carried out for the other four crimes as well. In the interest of space, the identification process results are shown in Table 5, while the estimated coefficients of the chosen models are given in Table 6. It should be noted that some of the ACs and PACs have relatively remote lags that seem significant. For example, in battery's PAC lags 2 to 10 are considered insignificant but lag 11 is significant. However, given the difficulty of finding a plausible reason why batteries eleven months ago impact present battery rates while those 2-10 months ago don't,[29] these lags are judged to be merely Type I errors. Indeed, given that the bands are 95% confidence intervals, one would expect the null of zero correlation to be falsely rejected 5% of the time, and thus it is not surprizing if out of forty lags a couple are misjudged. Thus, these lags are not included in the max ARMA considered.

The models suggested by the BIC are interesting for several reasons. For one, they demonstrate that every crime appears to show at least some serial correlation, as no ARMA(0,0) models were selected. For each of three ARMA(1,0) models chosen (for vehicle theft, arson, and battery), the AR(1) term was positive, indicating that the crime's frequency showed persistence. This is as expected, given studies previously noted that indicate significant persistence in crime frequencies. One can posit several reasons why this might be the case. Perhaps high rates of one of these crimes in the previous month made the crime appear less difficult, or more socially acceptable, changing the cost function facing the

[29] It is of course possible that some event occurs eleven months after a battery that causes recidivism. For example, perhaps battery perpetrators are typically released from prison after eleven months. However, no evidence of this sort of event could be found, and thus a Type I error seems more plausible. The same goes for the seventh lag in arson's AC and PAC, and other such outliers.

Table 5: Univariate Identification Results

Crime	Max ARMA considered based on AC and PAC	ARMA model chosen based on BIC	Q-test prob. value of residuals from minimum BIC model[30]
Burglary	ARMA(3,5)	ARMA(1,1)	0.9926
Robbery	ARMA(7,9)	ARMA(2,4)	0.9689
Vehicle Theft	ARMA(1,6)	ARMA(1,0)	0.9187
Arson	ARMA(3,3)	ARMA(1,0)	0.3394
Battery	ARMA(1,3)	ARMA(1,0)	0.8628

Source: Author's calculations based on CLEAR database.

Table 6: Univariate Model Coefficient Estimates

	Burglary	Robbery	Vehicle Theft	Arson	Battery
Constant	-10.58728	-.76422	1.80356	-0.00441	-0.84682
AR(1)	0.885***	1.63472***	0.79748	0.20135**	0.45927***
AR(2)	N/A	-0.70649***	N/A	N/A	N/A
MA(1)	-0.47321***	-1.20529	N/A	N/A	N/A
MA(2)	N/A	0.18864	N/A	N/A	N/A
MA(3)	N/A	-0.11976	N/A	N/A	N/A
MA(4)	N/A	0.41779	N/A	N/A	N/A

* significant at 10%, ** significant at 5%, *** significant at 1%[31]

Source: Author's calculations based on CLEAR database.

potential criminal. Or perhaps both periods are merely correlated with a third variable, such as income, as will be examined later in this paper. Whatever the case, it is clear that if one of these three crimes has a high frequency in month p, than month p + 1 will likely also have an above average frequency. For the other two crimes – burglary and robbery – the interpretation of their models is more difficult given the presence of more than one ARMA term. However,

[30] All p-values shown are for when 40 lags are included. However, all were also tested at 10 lags, 20 lags, and 100 lags, and for all the null hypothesis could not be rejected with 90% confidence.
[31] Based on standard errors calculated using the outer product of the gradient (OPG) method, the default method for estimating standard errors in Stata following the *arima* command.

the mere presence of meaningful lag terms again indicates that previous periods likely aid in the prediction of future crime. This claim will be examined when forecasts are made for each model.

Also of note is that – with the exception of robbery – each chosen model contains lags from at most a single period ago. This suggests that crimes occurring more than a month ago have little to no impact on crimes committed in the present. Why robbery seems to be the exception is an open question. Perhaps it is more heavily influenced by the economic cycle, causing more pronounced serial correlations.

Multivariate Analysis of Crimes

Having shown that each of the crimes examined appears to be at least somewhat correlated with own past values, the next step is to examine whether the crimes are in any way correlated with each other's past values. In particular, this paper is interested in whether past values of other crimes provide additional information that own past values do not. In other words, can a multivariate forecast improve on a univariate one? As stated previously, there are several reasons to believe it might. It is plausible that the chance of getting caught would change in similar ways across crime types, perhaps due to changes in police funding and technology.[32] As well, the opportunity costs of committing each type of crime may also move in similar directions, as forgone wages would presumably be closely related across crime types.

[32] That being said, changes in policing technology could also have a disproportionate effect on certain types of crimes and little effect on others, as described previously in relation to vehicle anti-theft technology.

Finally, changing social norms would likely effect different types of crimes in similar ways, and as the broken window theory hypothesises, there may be a feedback effect as well. Thus, if one crime responded to changes in economic, social, or policing norms more quickly than others, it could help predict changes in other crimes. As well, applying the logic of the broken window theory, multivariate forecasts may be optimal if a change in the prevalence of one type of crime led to shifts in the social acceptability of all crimes.

In order to judge if any of the types of crimes are useful in forecasting other types, Granger causality is tested for. Variable A "Granger causes" variable B with respect to information set S (where S contains A and B) if variable B is more accurately forecasted using the information set S as opposed to information set (S – A).[33] A method frequently used to test for Granger causality consists of constructing a vector autoregression (VAR) model using S, then testing the null that all coefficients on the lags of A in the equation for B are jointly equal to zero.[34] If this null is rejected, then A appears to Granger cause B, as its lags seem useful in explaining B.

Thus, VARs were constructed using all five studied crime types.[35] The method typically employed to determine the optimal lag length of a VAR is again to test down from a max lag length with the goal of minimizing some information criterion, such as the AIC[36] or BIC.[37] A max

[33] Granger (1969).
[34] Hacker and Hatemi-J (2006).
[35] As in the univariate section, the models in this section are constructed using the deseasonalized and detrended residuals of each crime frequency (with the exception of burglary which is only deseasonalized due to its judged lack of trend), as determined previously.
[36] The AIC, or Akaike information criterion, functions similarly to the BIC, but has the following formula: $AIC = 2k - (2) \cdot \ln(L)$, where again the L is the maximized likelihood value, and k is the amount of parameters being estimated.

lag of four was chosen as it seemed unlikely that the crimes would show significant correlation with other types of crime committed more than four months in the future, particularly in light of the fact that most of the types of crimes did not appear to show significant correlation with themselves more than one month in the future, with the exception of robbery which was correlated with itself four months in the future. It should be noted that this method assumes that each equation in the VAR has the same optimal amount of lags. As Table 7 shows, the BIC is minimized in the VAR(1) model, while the AIC is minimized in the VAR(4) model. Considering the ambiguity suggested in Table 7 over which model to use, Granger causality is tested for using both the VAR(1) and VAR(4) models, with the results shown in Table 8.[38] Indeed, given the univariate section showed large differences in the optimal lag length for different types of crimes, the assumption stated previously of identical lag lengths for each VAR equation may be too strong, and thus by examining both models should help to protect against this assumption biasing the results.

Table 7: VAR Selection Criterion

Model	AIC	BIC
VAR(1)	8116.041	8204.926
VAR(2)	8063.66	8226.231
VAR(3)	8031.993	8267.894
VAR(4)	7982.539	8291.411

Source: Author's calculations from CLEAR database.

[37] A good description of using selection criterion to determine VAR lag length can be found in Lutkepohl (1991) chapter four.
[38] The VAR(1) and VAR(4) equations used both to test for Granger causality and to generate Impulse Response Functions can be found in Appendix III.

Table 8: Tests for Granger Causality

Equation	Excluded	VAR(1) p-value	VAR(4) p-value
Burglary	Robbery	0.046	0.054
	Vehicle Theft	0.848	0.944
	Arson	0.386	0.738
	Battery	0.142	0.706
Robbery	Burglary	0.652	0.612
	Vehicle Theft	0.807	0.938
	Arson	0.074	0.341
	Battery	0.613	0.320
Vehicle Theft	Burglary	0.877	0.455
	Robbery	0.420	0.008
	Arson	0.191	0.000
	Battery	0.765	0.057
Arson	Burglary	0.224	0.245
	Robbery	0.021	0.006
	Vehicle Theft	0.486	0.185
	Battery	0.359	0.003
Battery	Burglary	0.002	0.052
	Robbery	0.143	0.143
	Vehicle Theft	0.831	0.160
	Arson	0.013	0.082

Source: Author's calculations from CLEAR database.

In the following section, the results of the Granger causality tests are examined and possible explanations for the results are hypothesised, with the aid of orthogonalized impulse response functions (IRFs), shown in Figures 4 and 5 (it should be noted that only IRFs suggested as pertinent by the Granger tests are included). Essentially, an IRF uses the VAR equations to forecast the response of variable B when a positive shock (or impulse) in variable A occurs.[39]

[39] Assessing the impact of a shock to a single variable becomes difficult when errors across variables are correlated. Thus, orthogonalized IRFs first orthogonalize the variance-covariance matrix using a process known as Cholesky

Thus, while Granger causality only tells us if variables are useful in forecasting each other, IRFs allow us some insight into the direction, duration, and extent of the relationship.

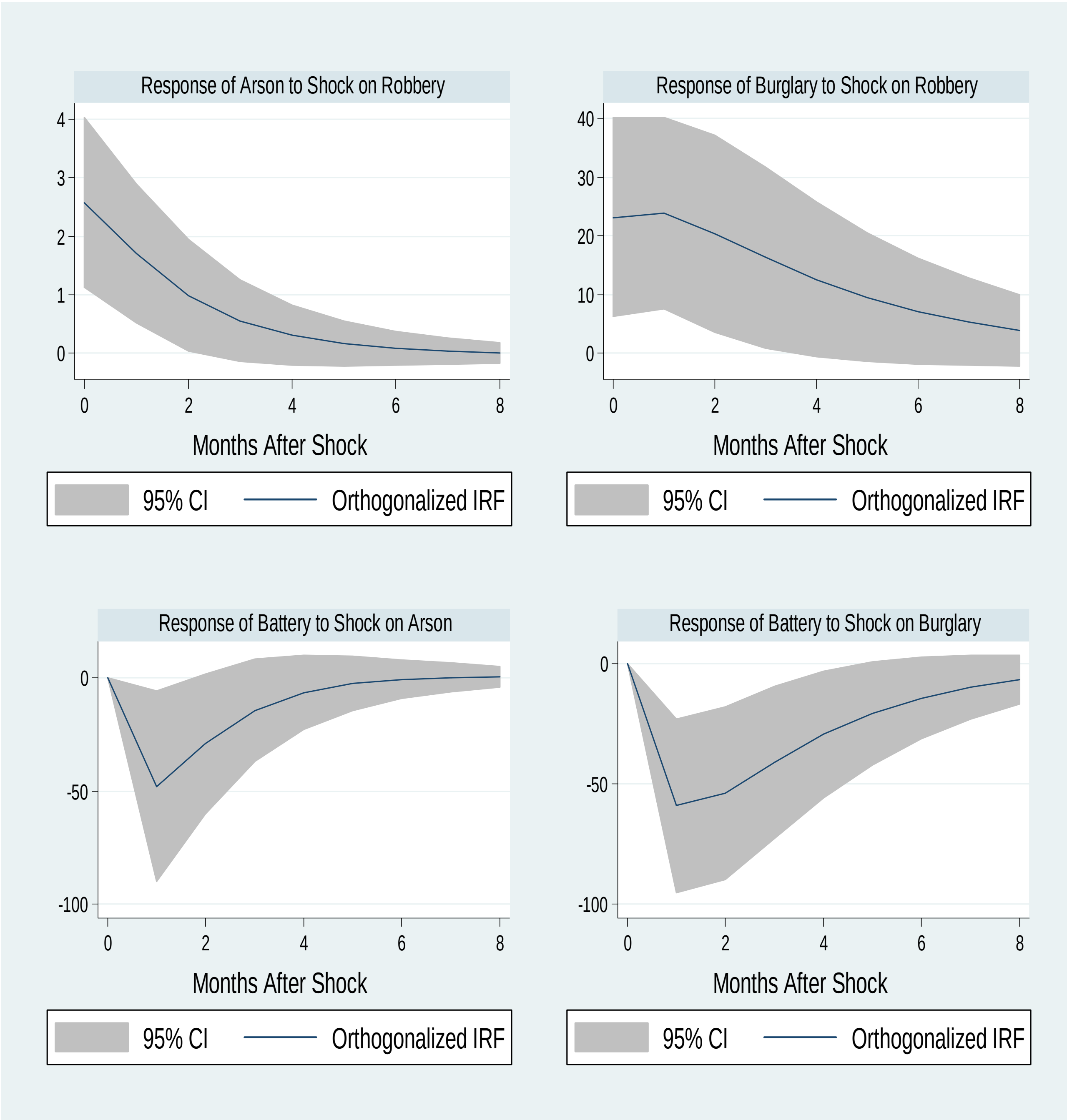
In the case of burglary, we can reject the null hypothesis that robbery does not Granger cause burglary in the VAR(1) case at the 5% level of significance, and can nearly reject it in the VAR(4) case. This indicates that past rates of robbery may be useful in forecasting future burglary rates. This result is perhaps not surprizing, as both burglary and robbery involve some form of larceny, and thus one would expect them to have similar determinants. As can be seen in Figure 4, the relevant IRF suggests that an increase in robberies is consistent with an increase in burglaries the following month, supporting the notion that both crimes are responding in a similar manner to a common cause. Concerning the robbery equation, we cannot reject the null for any of the other crimes, suggesting that past frequencies of other crimes would not be useful in forecasting robberies. The fact that robbery appears useful for predicting burglary but burglary is not useful for predicting robbery suggests that robbery may respond more quickly to changes in shared determinants.

For vehicle theft, while tests using the VAR(1) model do not indicate Granger causality by any of the other crimes, the VAR(4) tests point to both robbery and arson being possibly useful in prediction. Again, the inclusion of robbery makes sense due to the shared larceny component of both crimes. It is perhaps interesting that burglary does not seem Granger causal, as one might consider it more closely related to vehicle theft than robbery, which unlike the other two has a violent component. This fact may again point to robbery responding more

---

decomposition in order to avoid this problem. A more thorough description of orthogonalized IRFs can be found in Swanson and Granger (1997).

Figure 4: VAR(1) Orthogonalized IRFs



Source: Author's calculations from CLEAR database.

Figure 5: VAR(4) Orthogonalized IRFs

Source: Author's calculations from CLEAR database.

quickly to changes in common influences, such as economic conditions, thus making it more useful for predicting other types of crime, even if it is not closely related to them in practice. As expected, the relevant IRF in Figure 5 does show an increase in vehicle theft the month following a positive shock to robbery. However, this increase is not significant at the 95% level of confidence. Interestingly, the same IRF shows a statistically significant decrease in vehicle thefts four months following an increase in robberies. This could possibly be due to a response by the police force or heightened caution by the general public following the jump in robbery. Then again, it may simply be a statistical coincidence. Perhaps even more surprizing is the suggested Granger causality of arson with regards to vehicle thefts, as at first glance these appear to be quite different types of crime. It is certainly possible that arson and vehicle thefts are both correlated with economic changes; however direct relations between arson and economic circumstances do not immediately jump to mind.[40] Perhaps a more plausible explanation is provided by the broken window theory: arson is by definition among the most visible crimes committed, and thus could reasonably be expected to play a larger role than other less conspicuous crime types in shifting social norms towards illegal activity. It should be noted, however, that no statistically significant changes in vehicle thefts are recorded following a positive shock to arson, according to the relevant IRF in Figure 5.

Moving on to the arson equation, we can reject the null that robbery does not Granger cause arson in both the VAR(1) and VAR(4) models. The fact that robbery yet again appears important further cements our previous hypothesis that robbery responds more quickly to

[40] One possible connection: perhaps dire economic circumstances would lead to an increase in insurance fraud through arson.

common influences than other types of crime, thus serving as a bellwether for crime in general. The notion that an increase in robbery foretells an increase in other crimes is further supported by the relevant IRFs in Figures 4 and 5: the VAR(1) IRF shows a significant increase in arson rates the month following an increase in robberies, while the VAR(4) IRF's only statistically significant observation is an increase in arsons two months after an increase in robberies.

Continuing the examination of the arson equation, battery also appears to Granger cause arson in the VAR(4) case. Interpreting this result is difficult. It is perhaps notable that the relevant IRF in Figure 5 does not display any significant changes to arson rates following a shock to battery rates, suggesting that the relationship is not particularly strong.

While one would perhaps not expect arson and battery to be closely related, it is interesting to note that, according to the VAR(1) model, arson also appears to Granger cause battery. Indeed, the relevant IRF in Figure 4 suggests that an increase in arson rates is followed a month later by a significant drop in battery occurrences. Perhaps this is another example of the populace and police force responding to an increase in a high profile crime type by heightening their caution and crime avoidance tendencies, making the likelihood of getting caught rise and thus causing a fall in battery rates. Finally, burglary also appears to Granger cause battery, at least in the VAR(1) case. Again, according to the relevant IRF in Figure 4, the relationship appears to be negative as an increase in burglaries appears to result in a significant drop in batteries the following month. It is perhaps interesting that battery is the only crime studied displaying a significant negative response to increases in multiple other crime types. This is again difficult to interpret, and may merit further study.

The Granger causality results suggest some modification to our multivariate models is in order. An obvious course of action would be to construct an equation for each crime type that contains lagged values of only itself and the crimes judged useful by the Granger tests. In order to justify this course of action, however, we must first test to see if the crimes judged unnecessary by the Granger tests are indeed jointly unnecessary. To do this, we simply extend the Granger tests of each equation: whereas previously each null tested was that the coefficients on the lagged values of a single crime type were jointly equal to zero, now the null tested is that the lagged values of all the crime types judged insignificant by the previous Granger tests are jointly equal to zero. In the case of burglary, for example, since the earlier Granger tests suggested arson, car theft, and battery as not individually Granger causal, we now test the null that the lags of arson, vehicle theft, and battery in the burglary equation are jointly equal to zero, or in other words jointly not Granger causal.[41] The results are shown in Table 9.

Table 9: Joint Granger Causality Tests

Equation	**Crimes Tested to be Jointly Zero**	**VAR(1) p-value**	**VAR(4) p-value**
Burglary	Arson, Vehicle theft, Battery	0.370	0.958
Robbery	Burglary, Vehicle theft, Arson, Battery	0.459	0.630
Vehicle theft	Burglary, Battery	0.939	0.140
Arson	Burglary, Vehicle theft	0.417	0.094
Battery	Robbery, Vehicle theft	0.314	0.090

Source: Author's calculations from CLEAR database.

[41] It should be noted that only variables judged not Granger causal in both the VAR(1) and VAR(4) models are included. Thus, in the case of arson, only vehicle theft and burglary are jointly tested, despite the fact that battery is judged insignificant in the VAR(1) model.

Univariate and Multivariate Forecast Comparison

As can be seen in Table 9, none of the tests rejected the null at the 5% significance level, indicating that we can safely drop the variables tested. Having observed which variables appear useful for forecasting according to Granger tests, the next step is to actually carry out forecasts and compare the prediction accuracy of the multivariate and univariate models. First, new multivariate models are calculated according to the results from the previous section. Erring on the side of caution, multivariate models with both one and four lags are determined for each crime. Additionally, in cases where more than one additional crime was judged to be useful by the Granger tests, an equation was generated for each of the possible combinations of the useful crimes. Next, one-step ahead out of sample forecasts were calculated for each crime for the months from December 2010 to December 2012 (the final 25 periods observed), using the optimal univariate model as determined earlier as well as using each of the multivariate models discussed immediately above. For each forecast, the mean square error is calculated and used as a means to compare forecasting accuracy. The results are shown in Table 10.

Examining Table 10, the multivariate forecasts display a mixed record. Concerning burglary, robbery, and arson, the simple univariate models developed earlier generate the most accurate forecasts.[42] However, the inclusion of other crime types generates more accurate forecasts in the cases of vehicle theft and battery. Specifically, the most accurate forecast of vehicle thefts was generated by the VAR(4) model that included lagged values of itself and

[42] Admittedly, in the case of robbery only a univariate forecast was considered. However, given the results of the Granger tests, there is little reason to suspect that any multivariate forecast would outperform the univariate.

arson, while the most accurate forecast of battery came from the VAR(1) model including lagged values of itself and arson. The pertinent equations suggested by these models when all months are considered, are presented in Tables 11 and 12.

Table 10: One-Step Ahead Crime Forecasts

Forecasted Crime	Type of Model: Variables Used	MSE of Forecast
Burglary	**Univariate: ARMA(1,1)**	**13235.34**
Burglary	VAR(1): Burglary, Robbery	16037.74
Burglary	VAR(4): Burglary, Robbery	15367.96
Robbery	**Univariate: ARMA(2,4)**	**3057.017**
Vehicle Theft	Univariate: ARMA(1,0)	10384.3
Vehicle Theft	VAR(1): Vehicle Theft, Robbery, Arson	10268.86
Vehicle Theft	VAR(1): Vehicle Theft, Robbery	9989.349
Vehicle Theft	VAR(1): Vehicle Theft, Arson	10509.04
Vehicle Theft	VAR(4): Vehicle Theft, Robbery, Arson	9100.775
Vehicle Theft	VAR(4): Vehicle Theft, Robbery	11110.83
Vehicle Theft	**VAR(4): Vehicle Theft, Arson**	**8532.959**
Arson	**Univariate: ARMA(1,0)**	**58.67483**
Arson	VAR(1): Arson, Robbery, Battery	68.33259
Arson	VAR(1): Arson, Robbery	64.18393
Arson	VAR(1): Arson, Battery	61.67862
Arson	VAR(4): Arson, Robbery, Battery	67.01779
Arson	VAR(4): Arson, Robbery	72.54915
Arson	VAR(4): Arson, Battery	58.95981
Battery	Univariate: ARMA(1,0)	77041.97
Battery	VAR(1): Battery, Burglary, Arson	80207.82
Battery	VAR(1): Battery, Burglary	86823.82
Battery	**VAR(1): Battery, Arson**	**73186.21**
Battery	VAR(4): Battery, Burglary, Arson	82830.17
Battery	VAR(4): Battery, Burglary	90829.41
Battery	VAR(4): Battery, Arson	75694.92

Lowest MSE for each crime in bold.
Source: Author’s calculations from CLEAR database.

Table 11: VAR(4) Equation for Vehicle Theft, Using Vehicle Theft and Arson

Variable	Lag Length	Estimated Coefficient
Constant	N/A	0.1287
Vehicle Theft	Lag 1	0.8490***
	Lag 2	-0.1042
	Lag 3	0.0168
	Lag 4	0.0650
Arson	Lag 1	-1.5269
	Lag 2	1.3538
	Lag 3	0.7011
	Lag 4	-2.4513***

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database.

Table 12: VAR(1) Equation for Battery, Using Battery and Arson

Variable	Lag Length	Estimated Coefficient
Constant	N/A	5.0066
Battery	Lag 1	0.4726***
Arson	Lag 1	-5.3989**

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database.

It is interesting and quite surprizing to note that arson is the only crime type that improved our forecasts of other crimes. Concerning battery, the coefficients shown in Table 12 appear to suggest that an increase in the arson rate actually foretells a drop in battery the following month.[43] This is consistent with the IRF examined earlier. Again, this paper's best guess as to why this is the case is that the high profile of arson occurrences causes significant shifts in either social attitudes towards crime – making it more or less socially acceptable – or in

[43] Given multiple lags, it is more difficult to interpret the effect of arson on vehicle theft, but the strongly negative first and fourth lags of arson shown in Table 11 seem to suggest a negative relationship between arson and vehicle theft as well.

public awareness of crime – perhaps increasing the likelihood of getting caught due to heightened vigilance. Whatever the reason, it appears that multivariate analysis can, in select circumstances, help predict crime.

Panel Data Analysis

Having thus far demonstrated that the frequency of each crime type is related to lagged values of itself, and occasionally to lagged values of other crime types as well, we next wish to determine if crime is related to changes in income as well. As a measure of income, community median household income is used. As stated previously, given a lack of monthly income data, each crime type is aggregated to the annual level for this portion of the analysis, and only years 2003-2012 are studied. However, partially to help make up for this reduction in the sample size, the crime data (as well as median income data) are now also broken down by location: the annual frequency of each crime type is recorded for each of Chicago's 77 community areas. The resulting panel data thus contains 770 observations (77 community areas X 10 annual periods).

To determine the influence of income, the community frequency of each crime type is regressed on the community median income variable. Additionally, while seasonality is no longer an issue due to the use of annual data, the downward trend of each crime must still be dealt with, so a trend variable is also included in each regression. The first method used for regression is pooled OLS.[44] The results are shown in Table 13.

Table 13: Pooled OLS Regression Results

Type	Income Coefficient	Trend Coefficient
Burglary	-0.0033458***	-0.0248169
Robbery	-0.0038669***	-0.0926451***
Vehicle Theft	-0.0034502***	-0.1546532***
Arson	-0.0001786***	-0.0092541***
Battery	-0.0209212***	-0.7556219***

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database and US census data.

As before, all crime types except burglary display significant negative trends. Of more interest are the income coefficients. Each crime type displays a statistically significant negative relation with income. This appears to suggest that as income rises crime falls, perhaps because of the increased opportunity cost of forgoing honest work. However, given that pooled OLS was used to obtain these results, caution is in order. For instance, it is possible that relative income is negatively related to crime, but absolute income does not. Then richer neighbourhoods would display lower crime rates, but changes in income over time would not necessarily have an effect. The significance levels reported are also suspect, as pooled OLS works under the assumption that the error terms are iid. However, this is likely not the case, as fixed (or at least very slow changing) community attributes likely play a role in the prevalence of

[44] As the name implies, in this method all the panel data observations are pooled and then OLS regression is applied. A more thorough analysis of pooled OLS is provided in Wooldridge's *Econometric Analysis of Cross Section and Panel Data* (2002), pg. 169-173.

crime.[45] Thus, the error term could be written as $e_{it} = n_i + u_{it}$, where i represents the community area, t represents the time period, and n_i represents the "community effect" that is unchanging over time. This means that when moving from period t to period t+1, n_i will remain the same for each individual, making e display autocorrelation. As a result, pooled OLS will not be efficient and estimates based on it may have suspect standard errors.

Thus, in order to further test the hypothesis that higher incomes do have a negative effect on crime, results are also calculated using the Within-Groups transformation.[46] Suppose

$$Y_{it} = \beta X_{it} + n_i + u_{it} \tag{1}$$

To obtain the Within-Groups transformation, first the community means of both the left and right sides are taken:

$$\bar{Y}_i = \beta \bar{X}_i + n_i + \bar{u}_i \tag{2}$$

Then the Within-Groups transformation is simply the difference between equations (1) and (2):

$$(Y_{it} - \bar{Y}_i) = \beta(X_{it} - \bar{X}_i) + (u_{it} - \bar{u}_i)$$

Notice that the community effect n_i is cancelled out, so it is no longer an issue. Further, it is now clear that we are regressing on changes from the community mean over time, so the community mean itself has no bearing on the regression. Thus we can be sure that the impact measured is related to changes in income over time, as opposed to simply the difference in

[45] Such attributes could include geography or average age, to name but two examples.
[46] A more in-depth description of the Within-Groups method than provided here can be found in Arellano (1987). A brief explanation of the method can also be found in Davidson and MacKinnon (2004), pg. 298-300.

income levels between communities. The results of the Within-Groups transformation are display in Table 14.

Table 14: Within-Groups Regression Results

Type	Income Coefficient	Trend Coefficient
Burglary	-0.0036916***	-0.0238333
Robbery	-0.0018555***	-0.0718076***
Vehicle Theft	-0.0073642***	-0.1772506***
Arson	-0.0001185	-0.0086344***
Battery	-0.0136398***	-0.6740209***

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database and US census data.

In general, the results are quite similar to the pooled OLS case, indicating that an increase in a community's median income level really is correlated with a drop in most crime types. The notable exception is arson: whereas before its income coefficient was judged to be significant at the 1% level, in the Within-Groups estimate its coefficient is not judged significant even at the 10% level. The reason for this is twofold. For one, the estimated standard error of the Within-Groups coefficient was much larger than that of the pooled OLS ($7.5X10^{-5}$ as opposed to $1.7X10^{-5}$), likely because the pooled OLS error was underestimated due to the autocorrelation discussed previously. Second, the coefficient itself was considerably smaller in the Within-Groups model ($1.185X10^{-4}$ as opposed to $-1.786X10^{-4}$), suggesting that a large part of the effect estimated by pooled OLS was correlated only with differences between richer and poorer communities as opposed to changes in income levels over time. In the end, however, it is perhaps to be expected that arson would show less correlation with income than other crime types, as it is less directly related to monetary gain than burglary, robbery, or car theft.

Conclusion

Several conclusions can be drawn from this paper. The first is a confirmation that all the crime types studied show significant persistence, allowing one to use past values of a crime type to help predict its frequency in the future. In a way, this conclusion justifies the use of crime prediction discussed in the introduction. The usefulness of arson in predicting future rates of both battery and vehicle theft has also been established, suggesting that multivariate crime prediction methods may in some cases be optimal. Finally, it has been shown that increases in median community income and falls in crime are significantly correlated across all crime types except for arson.

In addition to these results, this paper also raises several questions that may merit future research. For one, the extreme persistence of robbery suggested by its univariate model relative to other crime types may warrant deeper analysis. Also, the uniqueness of arson, both in its usefulness in predicting other variables and in its lack of relation to income, is still largely unexplained. A better understanding of its determinates and how it relates to other types of crime could enable one to develop improved crime prediction models.

References

American Community Survey, United States Census Bureau. Accessed June 10, 2013 at http://www.census.gov/acs/www/

Arellano, Manuel. "PRACTITIONERS'CORNER: Computing Robust Standard Errors for Within-groups Estimators." *Oxford bulletin of Economics and Statistics*, Vol. 49 (1987): 431-434.

Ayres, Ian, and Steven D. Levitt. "Measuring Positive Externalities from Unobservable Victim Precaution: an Empirical Analysis of Lojack." *The Quarterly Journal of Economics*, Vol. 113 (1998): 43-77.

Bagshaw, Michael L. "Extension of Granger Causality in Multivariate Time Series Models." Research Department, Federal Reserve Bank of Cleveland, 1983.

Becker, Gary S. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, Vol. 76 (1968): 169-217.

Box, George E., and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day, 1970.

Cantor, David, and Kenneth C. Land. "Unemployment and Crime Rates in the Post-World War II United States: A Theoretical and Empirical Analysis." *American Sociological Review* (1985): 317-332.

Case, Anne C., and Lawrence F. Katz. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." National Bureau of Economic Research, Working paper 3705, 1991.

Cicero, Marcus Tullius. "Treatise On Laws: Book III [51 BC]," in *The Political Works of Marcus Tullius Cicero: Comprising his Treatise on the Commonwealth; and his Treatise on the Laws*. Translated by Francis Barham. London: Edmund Spettigue, 1841-42.

"City of Chicago Community Areas." Institute for Housing Studies, DePaul University. Accessed June 10, 2013 at http://www.housingstudies.org/page/city-chicago-community-areas/

"City of Chicago Data Portal series on crime." Chicago Police Department, CLEAR. Accessed June 8, 2013 at https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2

"Consumer Price Index: All Urban Consumers (CPI-U) U.S. City Average." Bureau of Labour Statistics, United States Department of Labour. Accessed Aug. 2, 2013 at ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

"Crime Type Categories: Definition & Description." Chicago Police Department, CLEAR. Accessed June 10, 2013 at http://gis.chicagopolice.org/clearmap_crime_sums /crime_types.html

Davidson, Russell, and James G. MacKinnon. *Econometric Theory and Methods*. Oxford: Oxford University Press, 2004.

Doyle, Joanne M., Ehsan Ahmed, and Robert N. Horn. "The Effects of Labor Markets and Income Inequality on Crime: Evidence from Panel Data." *Southern Economic Journal* (1999): 717-738.

Ehrlich, Isaac. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *American Economic Review*, Vol. 65 (1975): 397-417.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. "Efficient tests for an autoregressive unit root." *Econometrica*, Vol. 64 (1996): 813-836.

Fajnzylber, Pablo, Daniel Lederman, and Norman A. Loayza. *Determinants of crime rates in Latin America and the world: an empirical assessment.* World Bank Publications. 1998.

Fajnzylber, Pablo, Daniel Lederman, and Norman A. Loayza. "What Causes Violent Crime?" *European Economic Review*, Vol. 46 (2002): 1323-1357.

Farrell, Graham, Andromachi Tseloni, and Nick Tilley. "The Effectiveness of Vehicle Security Devices and Their Role in the Crime Drop." *Criminology and Criminal Justice*, Vol. 11 (2011): 21-35.

Freeman, Richard B. "Who Escapes? The Relationship of Churchgoing and Other Background Factors to the Socioeconomic Performance of Black Male Youths from Inner-City Tracts." In *The Black Youth Employment Crisis,* edited by Richard B. Freeman and Harry J. Holzer, 353-376. Chicago: The University of Chicago Press, 1986.

Gallup. "Crime: Gallup Historical Trends." Accessed June 8, 2013. http://www.gallup.com/poll/1603/crime.aspx#1.

Granger, Clive W.J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica*, (1969): 424-438.

Hacker, R. Scott, and Abdulnasser Hatemi-J. "Tests for Causality Between Integrated Variables using Asymptotic and Bootstrap Distributions: Theory and Application." *Applied Economics*, Vol. 38 (2006): 1489-1500.

Helland, Erik, and Alexander Tabarrok. "Does Three Strikes Deter? A Nonparametric Estimation." *Journal of Human Resources*, Vol. 42 (2007): 309-330.

Henry, Vincent E. *The COMPSTAT Paradigm: Management and Accountability in Policing, Business and the Public Sector*. Flushing, NY: Looseleaf Law Publications, 2003.

Kelling, George L., and James Q. Wilson. "Broken Windows." *Atlantic Monthly*, March 1982, 29-38.

Levitt, Steven D. "Understanding Why Crime Fell in the 1990s: Four factors that Explain the Decline and Six that Do Not." *Journal of Economic Perspectives*, Vol. 18 (2004): 163-190

Ljung, Greta M., and George Box. "On a Measure of Lack of Fit in Time Series Models." *Biometrika*, Vol. 65 (1978): 297-303.

Lutkepohl, H. *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag, 1991.

Maslow, Abraham. "A Theory of Human Motivation." *Psychological Review,* Vol. 50 (1943): 370-396.

Myers, Samuel L. "Estimating the Economic Model of Crime: Employment Versus Punishment Effects." *Quarterly Journal of Economics*, Vol. 98 (1983): 157-166.

National Research Council. *Using the American Community Survey: Benefits and Challenges*, edited by Constance F. Citro and Graham Kalton. Washington, DC: The National Academies Press, 2007.

"News Release: GDP by State." United States Department of Commerce, Bureau of Economic Analysis. Accessed Aug. 15, 2013, at http://www.bea.gov/newsreleases/regional/gdp_state/gsp_newsrelease.htm

Paral, Rob. "Chicago Census Data." Rob Paral and Associates. Accessed June 10, 2013 at https://docs.google.com/spreadsheet/pub?key=0AtvdVdII0XSCdE9NY3lEek5jTEFzdGF2R2k4TWRqOEE&gid=0

Paral, Rob. "Median Household Income in Chicago Community Areas 2006-2010." Rob Paral and Associates. Accessed June 10, 2013 at https://www.google.com/fusiontables/DataSource?docid=1DZWxPHHoRTVAs8Gnqz3y7xvGnuKYlbm2Cuf5Scw#rows:id=1

Perez, Douglas Werner, and Michael Barkhurst. *Paradoxes of Leadership in Police Management*. Clifton Park, NY: Delmar Cengage Learning, 2012.

"Police-Reported Crime Rates, Canada, 1962-2011." Statistics Canada, Canadian Centre for Justice Statistics, Uniform Crime Reporting Survey. Accessed Aug. 2, 2013, at http://www.statcan.gc.ca/pub/85-002-x/2012001/article/11692/c-g/desc/desc01-eng.htm

Schwarz, Gideon. "Estimating the Dimension of a Model." *The Annals of Statistics*, Vol. 6 (1978): 461-464.

Schwert, G. W. "Tests for Unit Roots: A Monte Carlo Investigation." *Journal of Business and Economic Statistics*, Vol. 7 (1989): 147–160.

Swanson, Norman R., and Clive W.J. Granger. "Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions." *Journal of the American Statistical Association* (1997): 357-367.

Trumbull, William N. "Estimations of the Economic Model of Crime Using Aggregate and Individual Level Data." *Southern Economic Journal* (1989): 423-439.

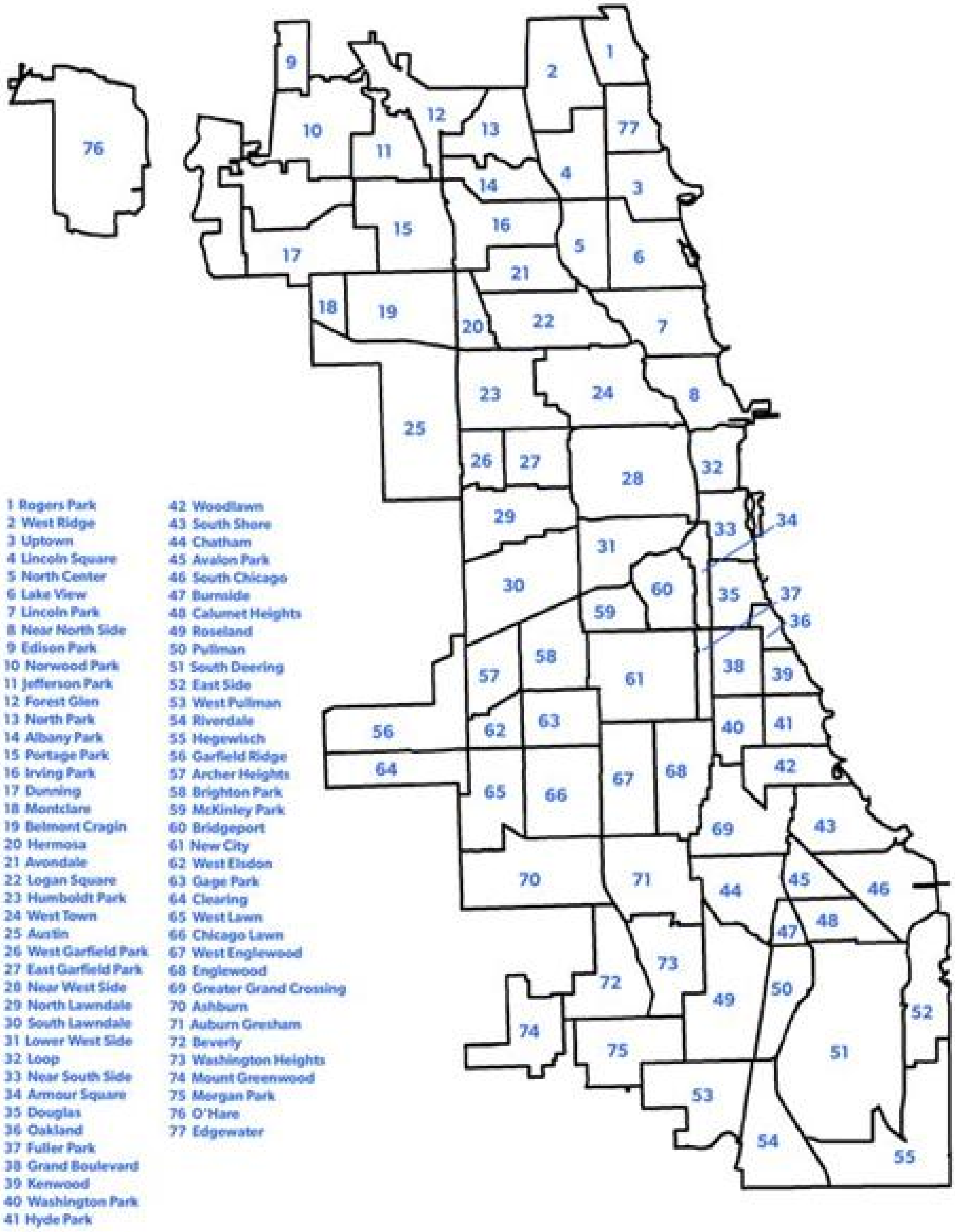
"Uniform Crime Reporting Data Online." Department of Justice. Accessed June 8, 2013 at http://www.ucrdatatool.gov/

Weisburd, David, Stephen D. Mastrofski, Rosann Greenspan, and James J. Willis. "The Growth of CompStat in American Policing." Police Foundation, April 2004.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

Yang, B., and D. Lester. "The Deterrent Effects of Executions: A Meta-Analysis Thirty Years After Ehrlich." *Journal of Criminal Justice*, Vol.36 (2008): 453-460.

Appendix I: Chicago Community Area Map



Source: "City of Chicago Community Areas," Institute for Housing Studies.

Appendix II: Determination of Annual Community Median Income Figures

The purpose of this appendix is to describe how annual median household income levels for each community area were obtained. Initial data collection only yielded three income figures for each community: the 2000 Census figure, the 2006-2010 ACS estimate, and the 2007-2011 ACS estimate. In order to estimate the annual figures from 2006-2011, the real annual growth rate of Chicago-wide median household income was calculated using the ACS one-year estimates[47], adjusted for inflation.[48] These growth rates were then used to estimate each year within the five-year estimates (FYEs), using the following method:

> Suppose the FYE was of years t, t+1, t+2, t+3, and t+4, and let median household income in period t be represented by x_t and the real growth rate from periods t to t+1 be g_t.[49] Then, given that the FYE is essentially an average of data collected in each of the five years,
>
> $5FYE = (x_t) + (x_{t+1}) + (x_{t+2}) + (x_{t+3}) + (x_{t+4})$
>
> $= (x_t) + (x_t)(g_t) + (x_t)(g_t)(g_{t+1}) + (x_t)(g_t)(g_{t+1})(g_{t+2}) + (x_t)(g_t)(g_{t+1})(g_{t+2})(g_{t+3})$
>
> $\Rightarrow x_t = 5FYI / [(g_t) + (g_t)(g_{t+1}) + (g_t)(g_{t+1})(g_{t+2}) + (g_t)(g_{t+1})(g_{t+2})(g_{t+3})]$
>
> Having calculated x_t, determining x for years t+1, t+2, t+3, and t+4 is relatively straightforward.

For years 2007, 2008, 2009, and 2010 – where two figures are obtained for each community due to the overlap of the five-year estimates – the average of the two estimates is taken.

[47] As stated previously, one-year estimates are available only for high population areas, such as Chicago as a whole.
[48] Using the United States Consumer Price Index for all urban consumers (CPI-U).
[49] We use real growth rates as the five-year estimates are already adjusted for inflation.

While annual figures have thus been obtained for years 2006-2011, figures for years 2003-2005 and 2012 must still be obtained. Unfortunately, given the relatively recent introduction of the ACS, one-year Chicago estimates are not available for years before 2005. Instead, we rely on the 2000 Census data for each community, and assume an identical annual growth rate (g) over the years from 1999[50] to 2006. Thus $x_{2006} = (x_{1999})(g^7)$, so using our previously obtained 2006 median income estimate as well as our 2000 Census figure, we are able to obtain a specific g for each community area. Using this g, we can then easily estimate annual median income figures for 2003, 2004, and 2005.

Not surprisingly, it was difficult to find Chicago growth data for 2012, given its proximity to the present. Thus, the annual GDP growth rate of Illinois in 2012 (1.9%)[51] was simply applied to the 2011 median household income figure for each community in order to obtain a 2012 estimate.

[50] Note that the 2000 Census median household income data concerns income earned in 1999.
[51] Obtained from "News Release: GDP by State," Bureau of Economic Analysis, United States Department of Commerce.

Appendix III: VAR(4) and VAR(1) Models Using All Crime Types

Table A3.1: VAR(4) Equations Containing All Five Crime Types

Variable	Lag Length	Coefficient in Burglary Equation	Coefficient in Robbery Equation	Coefficient in Vehicle Theft Equation	Coefficient in Arson Equation	Coefficient in Battery Equation
Constant	N/A	-2.7219	-0.0771	-1.3957	0.1241	8.5829
Burglary	Lag 1	0.3989***	0.0150	-0.0827	-0.0063	-0.5901***
	Lag 2	0.1963**	-0.0100	0.1167	-0.0121	0.1049
	Lag 3	0.1992**	0.0813	0.0022	0.0039	-0.0754
	Lag 4	-0.0408	-0.0641	0.0824	0.0004	0.1308
Robbery	Lag 1	0.2882**	0.6314***	0.1991	0.0118	0.2930
	Lag 2	-0.0062	0.0855	-0.4317***	0.0323**	0.4462
	Lag 3	-0.3978**	-0.2250**	-0.1796	-0.0139	-0.2224
	Lag 4	0.2583	0.2304**	0.1021	0.0136	-0.6590*
Vehicle Theft	Lag 1	-0.0315	0.0371	0.8401***	0.0099	0.0909
	Lag 2	-0.0226	-0.0551	-0.1393	-0.0199**	-0.0630
	Lag 3	0.0181	0.0311	0.0438	0.0200**	-0.0995
	Lag 4	0.0395	-0.0188	0.0637	-0.0069	-0.2175
Arson	Lag 1	0.0025	-0.4688	-1.7805*	0.1903**	-5.8625**
	Lag 2	-1.2505	-0.8454	3.1983***	-0.1179	-0.0055
	Lag 3	1.0567	0.6888	0.7514	0.0656	3.4499
	Lag 4	-0.1341	.4013	-2.7159***	0.1028	.4377
Battery	Lag 1	-0.0210	-0.0039	0.0839**	-0.0045	0.3874***
	Lag 2	0.0124	0.0460**	-0.0837**	0.0065**	-0.0009
	Lag 3	-0.0309	-0.0133	0.0244	0.0000	0.0767
	Lag 4	0.04171	-0.0771	-0.0291	-0.0089***	-0.0689

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database.

Table A3.2: VAR(1) Equations Containing All Crime Types

Variable	Lag Length	Coefficient in Burglary Equation	Coefficient in Robbery Equation	Coefficient in Vehicle Theft Equation	Coefficient in Arson Equation	Coefficient in Battery Equation
Constant	N/A	-1.5265	0.1574	0.7252	0.0831	6.0974
Burglary	Lag 1	0.5427***	0.0199	-0.0115	-0.0077	-0.5605***
Robbery	Lag 1	0.2445**	0.6593***	-0.0987	0.0241**	0.4340
Vehicle Theft	Lag 1	-0.0096	-0.0073	0.7999***	0.0030	-0.0258
Arson	Lag 1	-0.8588	-1.0552*	-1.2971	0.1511*	-5.9767**
Battery	Lag 1	-0.0436	0.0090	0.0089	-0.0023	0.4396***

* significant at 10%, ** significant at 5%, *** significant at 1%
Source: Author's calculations from CLEAR database.