UNITED STATES MILITARY ACADEMY

ANALYZING IMPACTS ON BURGLARY FREQUENCY

MA478

SECTION H2

COL. CLARK

BY

CADET AIMEE ROHAN'25, CO I3

WEST POINT, NEW YORK

7 MAY 2024

AR_ MY DOCUMENT IDENTIFIES ALL SOURCES USED AND ASSISTANCE RECEIVED IN COMPLETING THIS ASSIGNMENT.

_ I DID NOT USE ANY SOURCES OR ASSISTANCE REQUIRING DOCUMENTATION IN COMPLETING THIS ASSIGNMENT.

SIGNATURE:

Chicago Burglary: Impacts of Socio-Economic and Temporal Factors

Aimee Rohan Ramirez

April 2024

Abstract

Research conducted at the city level has demonstrated that burglary rates tend to fluctuate based on factors such as wealth, unemployment, population, season, and year. This study, focusing on data at the census-block level, seeks to deepen our comprehension of how temporal and socio-economic variables influence burglary frequency in Chicago. To achieve this, a two-step methodology is employed. First, we take into consideration and explore which factors–unemployment, wealth, time, and year have impacts on burglary rates. Second, once we use models such as Poisson, Zero-Inflated Poisson model (ZIP), Generalized Linear Mixed-Effects model (GLMM), and and an Auto-Regressive GLMM to explore taking into consideration whether there are also unique random effects placed on month, year, or census block that define burglary rate in Chicago. Our models yielded mixed results, with factors such as unemployment and year, as well as the inclusion of a random effect of months, showing varying impacts. However, the performance of the models, indicated by poor AIC values or failure to pass goodness-of-fit tests, complicates drawing conclusions regarding the influence of temporal and socioeconomic characteristics. This is further compounded by the inconsistent findings observed.

Key Words: Crime trends, Chicago, Temporal Impacts, Generalized Linear Models, Random Effects

1 Introduction

Crime is present everywhere. Not only that, burglary is present in every state and country everywhere. Moreover, federal officials have tried to reduce the onset of this crime through the introduction of bills that fund law enforcement officers with the tools and protection they need in order to address burglary rates ([gov24]). As a result of these policy changes, a wide array of research questions can be considered. For our purpose, we focus on socio-economic factors to which policies have associated with lowering burglary frequency in Chicago's Census Block groups. At the same time most studies that do exist consider additional crimes with also possibly considering seasonal patterns or additional characteristic factors such as race, impact of COVID, rates of disease, or spatial characteristics ([CFAP20, DSD23, Fre18]). However, these studies have been unable to separate monthly and yearly impacts while incorporating whether each census block has a unique random effect to add. This current study pertains to this issue. Herein, we use data from 255 Chicago Census Blocks from the years 2010-2015 to examine whether socio-economic and temporal aspects are correlated with raising or decreasing burglary frequency.

2 Literature Review

Crime remains an enduring challenge across the United States. Given its persistent nature, criminologists have delved into the realm of policy, seeking to discern the extent to which various policies have shaped the prevalence of different types of crime [LJ10]. While research on policy impact is imperative in addressing crime frequency, delving deeper into understanding the dynamics between micro and macro-level patterns in society is essential for comprehending changes in the frequency and types of crime. Moreover, research can indicate that crime does not occur randomly across space and time [Wei15, MLP12, CF79]. Events can rather follow a pattern based on social, economic, and demographic conditions of the environment. Despite trying to account for the pattern in Burglary crimes by adding wealth, count of young men, and unemployment into our study, we face the limitation shared by numerous studies, [CFAP20, Fre18, LJ10], being that there is a gap present between actual and reported crimes. It is also why we used a ZIP model in order to explain that there are possibly many non-reports of burglary in a given month due to the data depending on the mechanism whether one reports or not.

Research as well indicates that crimes do not behave the same way across different areas in Chicago [CFAP20]. Campedelli's adjusted for this complexity through the addition of Bayesian models to add a spatial perspective yet also adding in various crime related, socio-economic, and health-demographic characteristics to see if there was a reduction in crimes due to COVID's impacts. We excluded Bayesian models from our analysis due to the absence of spatial characteristics. To address this gap, we introduced an error term unique to each Census Block, assuming that each block contributes uniquely to the burglary count. Our current study attempts to disentangle some of the socio-economic characteristics in Chicago before COVID while also considering the unique random effects each census block has in affecting the frequency of burglaries.

It is commonly accepted that most types of crime show seasonal fluctuations. Seasonality here refers to a cyclical pattern that repeats at regular intervals [DSD23]. It is important to capture this cyclical effect as most studies since the 20th century have found that the highest aggregated offenses of burglary or violent crimes have occurred in the summer [CF79, MLP12, DSD23]. Two components that arise when discussing seasonality is whether criminal activity arises due to changes in social behavior or temperature.To address this question, McDowall's study in *Seasonal Cycles in Crime, and their Variability* concluded that monthly temperature variations alone could not fully explain the influence on crime [MLP12]. Social behavior and activities emerged as additional factors to consider. In similarity to Delgado's study and McDowall's conclusion, we have made the decision to not include average temperature to our study as we reasoned that temperature overall is heavily correlated with the month of the year and our assumption was that our geographical units, being Census Blocks in Chicago, carry a similar temperature to one another. Where we have differed in Delgado's study in including seasonality, was that we did not use Colwell's metrics to capture the different aspects of seasonality in a given geographical unit as we did not build a supervised learning algorithm to predict the level of crime based on the time unit. Our purpose was different as we wanted to infer in what way does month or year given our time frame affect the frequency of burglaries in the Census Blocks of Chicago. Moreover, our approach was similar to McDowall's by adding month as a factor in our analysis in order to capture the cyclical effect seasonality has.

3 Methodology

To analyze the amount of burglaries in Chicago, we compile a dataset from raw data retrieved from a provided link that includes burglary counts for each month from 552 distinct Census block groups within Chicago measured from 2010-2015. The GitHub raw data also features other datasets covering various socio-economic factors, including an adjacency matrix indicating neighboring counties, population statistics, unemployment rates, wealth distribution, and demographic data such as the count of young males in each location. With this data we planned to use population stats, unemployment rates, wealth distribution, and the demographic data to explore our purpose.

The objectives set out in this work covers two aspects. It would be interesting by the end of this report to have been able to answer the following questions raised by burglary in the Census Blocks of Chicago from 2010-2015:

1. Do the characteristics of unemployment, wealth, or count of young males impact burglary count?
2. Does Seasonality impact the frequency of burglaries in the Census Blocks of Chicago? Does taking into consideration the year as a time variable improve our models?

3. Should we consider month solely as a covariate or does our models benefit by adding a random effect? Is there a random effect based on the uniqueness of a Census Block?

This study takes into consideration the Routine Activities theory where it offers a comprehensive approach to seasonality where it can explain patterns in property crimes. The theory calls to attention how temporal trends help structure individual behaviors in which affect the risk of victimization. Overall the foundation of this theory is based on the idea that environmental conditions create an opportunity to commit a crime [MLP12]. Moreover, by using the routine activities approach, and by some studies concluding that average temperature is highly correlated with seasonality [MLP12, CF79, DSD23], we opted to not include average temperature in our models.

Questions one and two are addressed by the exploration of the data and by building, validating and comparing different statistical models. Additionally, by adding a month as a factor in our models we are able to determine which months are the most impactful on burglary frequency in the Census Block of Chicago.

Question three is answered through the building and analysis of a generalized linear mixed model (GLMM) and a auto-regressive Poisson GLMM. The performance of the GLMM strengthens the argument that the uniqueness of the Census Block and month explain the burglary Frequency.Through the auto-regressive model, we can analyze the impact the last month has on explaining the month after in terms of burglary frequency.

4 Experimentation and Results

4.1 Data Preparation

We loaded the several CSV files hosted on Github.These datasets include crime counts, wealth by Census Block Group, population by Census Block Group, percentage unemployed by Census Block Group, and the number of young males by Census Block Group. Each dataset is loaded into a data frame format.

We then conducted some reshaping on our crime count dataframe to have a separate row for each month and year combination while also assigning a block ID to indicate the Census Block Group to which the crime count belongs. We then merged the various datasets together into a single dataframe based on the Census Block ID shared by all datasets.

For handling date formatting, we extracted the month and year from a column that had them together and convert them into separate numeric columns. Moreover, we standardized the population variable by subtracting the mean and dividing by the standard deviation. We also adjusted the year variable to to start from 2010. This adjustment was necessary to fit the data into a specific analytical framework for our models.

Overall, we used two different dataframes for our models. One dataframe is a replica of the above only that it has the month as a categorical factor in order to encapsulate the cyclical pattern of seasonality. The second dataframe has month as a quantitative factor where there is a linearity assumption in order to capture overall, how months play an impact in burglary frequency. This dataframe matters for models that want to capture how much total of the month before explains the months after burglary frequency.

4.2 Data Exploration

The mean number of burglaries per census block was 1.203 properties, with a standard deviation of 1.481, indicating variability in burglary rates across blocks. In Figure 4, each census block exhibits an average of approximately 1 burglary, with occasional outliers showing approximately 3 to 4 burglaries.

When conducting a uni-variate analysis on burglary frequency, we see that the number that is reported most often in Figure 1 is zero. This is when we considered that this could be due to the data depending on a mechanism where one reports or not. Due to there being a possible mechanism present, it presents the likelihood of a ZIP model fitting the data well due to the high number of zero's present.

In the visual analysis burglary frequency by month, Figure 2, a clear pattern emerges: burglary rates decline during colder months (from months 11 to 02). This trend suggests that individuals may be more inclined to stay indoors during the colder weather, particularly during major holidays, aligning with the principles of Routine Activities Theory. It's essential to note that temperature and seasonality are closely correlated [DSD23, MLP12]. Consequently, we opted not to include average monthly temperatures when we built our models.

When considering if there were any temporal variables impacting burglary frequency, we looked at the burglary count across the years 2010 to 2015. In Figure 3, we see a decline throughout the years. What is difficult is determining why there is a decline throughout the years. This could be due to many factors such as the mechanism to report being made more difficult, more police officers around, or if people did not want to report. This is taken into consideration as a limit in our future models as there could be a gap between actual and reported burglary crimes.

In conducting bivariate analysis, as depicted in Figure 5, there is a slight positive correlation between higher wealth ratings of census blocks and increased average burglary counts within those blocks. While this finding is based on visual analysis and no formal statistical tests were conducted, the trend suggests that Census blocks with a higher wealth rating do appear to be at a higher risk for burglary.

In our visual analysis depicted in Figure 6, we clearly observe a strong positive association between the count of young men and wealth rating, which signifies the presence of multicollinearity. To address this issue, the count of young men variable was deliberately excluded from our models in order to ensure stability and minimize collinear influences. Consequently, it is acknowledged that our models may not fully disentangle the intertwined effects of wealth rating and counts of young men per Census Block on burglary frequency.

4.3 Model Building and Assessment

Model 1: Poisson Generalized Linear Model (GLM)

With using this model, we do acknowledge that we are assuming that there is no over-dispersion present in the data given. Meaning, that our mean and variance are equal. Based on our exploratory data analysis, variables such as wealth, population, and unemployment rate were found to have an slight association with burglary frequency, thus included in our initial model.

Additionally here in this model we added population as an offset in our model in order to account for differences in Census Block density. That is, we are predicting burglary count per population instead of solely burglary count. Moreover, we included month in our model as a categorical factor in order to capture the cyclical pattern into our model.

$$
\begin{aligned}
y_{ijk} &= \text{count of burglary for block } i \text{ in year } j \text{ and month } k \\
y_{ijk} &\sim \text{Pois}(\lambda_{ij}) \\
\eta_{ijk} &= \log(\lambda_{ijk}) = \beta_0 + \beta_1 \text{Wealth}_i + \beta_2 \text{Unemployment}_i + \beta_3 \text{Year}_j \\
&\quad + \gamma_1 \text{January} + \ldots + \gamma_{12} \text{December} + \log(\text{Population}_i)
\end{aligned} \tag{1}
$$

Wealth, unemployment, and population for each Census Block i are represented as a fixed effect ; γ_1 through γ_{12} are the coefficients set for the months as fixed effects.

The incorporation of fixed effects addresses the consistent variations in crime rates among different census blocks, while the inclusion of a linear year trend accommodates systematic changes in burglary frequency over time. Additionally, treating the month variable as categorical enables the assessment of seasonal fluctuations in burglary frequency, acknowledging that certain months may exert more influence than others.

Results: (reference Table 1 for the coefficients for this model)

The model through Table 1 show a seasonal pattern that we noted earlier during our data exploration. All the months were shown to be deemed as significant effects in determining burglary frequency. Burglary tends to decrease in January, February, and March. In comparison to our reference month,January, the expected number of crimes in August increases by a factor of 1.377.

Unemployment throughout the years 2010 to 2015, were shown to increase burglary frequency by a factor of 1.59. To add on, for each one unit increase in wealth, the expected count of burglary increases by a factor of 1.18.

It's crucial to acknowledge that this model fails the Goodness of Fit test, indicating that it does not sufficiently capture the variance in the data. Consequently, any inferences drawn from the results of this model may be questionable at present. This undermines the validity of any conclusions or recommendations based on its findings. Overall, this underscores the need to explore alternative models that provide a better fit to the data.

Model 2: Zero Inflated Poisson Model

In Figure 1, it is evident that a considerable number of reported burglaries from 2010 to 2015 are zero. This observation suggests a prevalence of zero-inflation in the data, possibly stemming from reporting mechanisms. To address this issue, we employed a Zero-Inflated Poisson (ZIP) model, which incorporates the same explanatory variables as the previous model. What we are assuming by using a ZIP model is independence. Meaning, that the presence or absence of excess zeros does not depend on the magnitude or frequency of non-zero counts. The ZIP model formulation is as follows:

$$
\begin{aligned}
&y_{ijk} \sim \begin{cases} 0 & \text{w/p } \phi_{ijk} \\ Pois(\lambda_{ijk}) & \text{w/p } (1-\phi_{ijk}) \end{cases} \\
&z_{ijk} \sim Bern(\phi_{ijk}) \\
&y|z_{ijk} = (\frac{e^{-\lambda_{ij}}\lambda^2}{y!})^{1-z_i} \\
&\eta_{ijk} = log(\lambda_{ijk}) = \beta_0 + \beta_1 Wealth_i + \beta_2 Unemployment_i + \beta_3 Year_j + \gamma_1 January + ... + log(Population_i)
\end{aligned} \tag{2}
$$

Wealth, unemployment, and population represent the values for each Census block i as a fixed effect; γ_1 through γ_{12} are the coefficients set for the months as fixed effects.

Having month as a categorical dummy variable in the ZIP model allows for a more flexible modeling of the seasonal variation in burglary frequency. By treating month categorically in this model allows for the estimation of separate effects for each month, providing insights into the seasonal patterns of burglary occurrence.

Results: (reference Table 2 for the coefficients for this model)

The intercept of the zero-inflated model coefficient being -1.7973 indicates that the log odds of observing zero burglaries is low. This suggests that there is a low probability of observing zero reports of burglaries. This observation aligns with the increased AIC value for the zero-inflated model compared to Model 1, as shown in Table 5. In the count model, both wealth and unemployment were deemed significant, with their p-values being less than 0.05. Higher values of these variables were associated with increased burglary counts.

The statistically significant coefficients associated with specific months in both the count and zero-inflation models underscore the nuanced seasonal patterns in burglary frequency. Notably, months like August and September demonstrate pronounced positive effects in the count model, implying heightened burglary activity during these periods. What stands out is also the conclusion in this model that winter months have lower burglary frequency.

Model 3: Generalized Linear Mixed-Effects Model (GLMM)

The proposed generalized linear mixed model accounts for the presence of unique random effects attributed to both census block and month variables. This approach acknowledges potential heterogeneity in burglary counts across different geographic areas (blocks) and temporal periods (months), providing a more nuanced understanding of the underlying data dynamics. What we did different in this model was that we had month listed as a quantitative variable in order to improve the interpretability of this model. Since month is a random effect in this model, using it as a quantitative variable allows the model to account for the continuous variation in burglary counts across different months more accurately.

However, it's essential to address certain assumptions inherent in this modeling framework. Firstly, the model assumes that the random effects follow a multivariate normal distribution with mean zero and variances $\sigma^2_{\text{block}}, \sigma^2_{\text{month}}$ for blocks and months. This assumption implies that the random effects capture unobserved variability in burglary counts not explained by the fixed effects. Additionally, the model assumes that the observed crime counts follow a Poisson distribution with a mean parameter λ_{ijk}.

It's worth noting that the model encountered convergence issues when attempting to include a standardized population offset. This suggests potential complexities in the data structure or model specification that need to be addressed or investigated further to ensure robust model estimation and inference.

$$
\begin{aligned}
&y_{ijk} = \text{burglary count for census block } i \text{ in year } j \text{ and month } k \\
&y_{ijk} \sim \text{Pois}(\lambda_{ijk}) \\
&\log(\lambda_{ijk}) = \eta_{ijk} = \beta_0 + \beta_1 \cdot \text{wealth\_d}_i + \beta_2 \cdot \text{un.emp}_i + \beta_3 Year_j + b_{0i} + b_{1k} \\
&\text{where} \\
&\quad y_{ijk} \text{ is the observed crime count,} \\
&\quad \lambda_{ijk} \text{ is the expected mean of the Poisson distribution,} \\
&\quad b_{0i} \sim \mathcal{MVN}(0, \sigma^2_{\text{block}}) \text{ represents the random effect for block } i, \\
&\quad b_{1k} \sim \mathcal{MVN}(0, \sigma^2_{\text{month}}) \text{ represents the random effect for month } k,
\end{aligned}
\tag{3}
$$

Matching the above models, Wealth, unemployment, represent the values for each Census block i as a fixed effect; year is interpreted as a quantitative impact based on year j.

Results: (reference Table 3 for the coefficients for this model)

The intercept coefficient represents the estimated log expected burglary count when all other predictors are zero. Here, it is statistically significant ($p < 0.001$), indicating that even when wealth, year, and unemployment rate are zero, there's still a non-zero expected burglary count.

The coefficient for wealth is also statistically significant ($p < 0.001$), suggesting that an increase in wealth rating is associated with a proportional increase in the expected burglary count, holding other variables constant.

The coefficient for unemployment is not statistically significant ($p = 0.33$), suggesting that there is insufficient evidence to conclude that unemployment rate has a significant effect on burglary counts after accounting for other predictors.

Despite this model exhibiting the lowest AIC value compared to the other models (refer to Table 5), it is important to note that this model incorporates random effects, which influence the structure of the model and the estimation of η. The absence of a standardized population offset in this model contributes to its uniqueness and makes it unsuitable to use AIC for direct comparison with other models.

Model 4: Auto Regressive Poisson GLMM

Our last proposed model aims to address the temporal aspect of the problem more comprehensively by introducing random effects due to Census Block and month, thus forming a generalized linear mixed model (GLMM). This approach acknowledges that the frequency of burglaries may vary not only across different years but also within specific blocks and months. In this model as well we captured month as a quantitative factor over categorical in order to capture how much a month prior explains the burglary frequency after. To capture potential correlations between consecutive years, we incorporate an auto-regressive correlation structure (AR(1)) on the year variable, represented by u_k. Our model can be written as follows:

$$
\begin{aligned}
&y_{ijk} = \text{count of burglary for block } i \text{ in year } j \text{ and month } k \\
&y_{ijk} \sim \text{Pois}(\lambda_{ijk}) \\
&\eta_{ijk} = \log(\lambda_{ijk}) = \beta_0 + \beta_1 \text{Wealth}_i + \beta_2 \text{Unemployment}_i + \beta_3 Year_j + u_k + \alpha_i \\
&u_k = \phi_{u_{k-1}} + \upsilon_k \\
&\upsilon_k \sim (0, \sigma_\upsilon^2) \\
&\alpha_i \sim N(0, \sigma_\alpha^2)
\end{aligned}
\tag{4}
$$

Where u_k represents AR(1) correlated errors associated with different years, σ_α^2 represents the variance of the random effect due to block, α_i, and $\Omega \sim N(0, \sigma_\Omega^2)$ represents the variance of the random effect due to month, Ω_j.Both random effects are assumed to follow a normal distribution with mean 0. This variance parameter captures the variability of the random effects among different census blocks. The variance in our model captures the variability in burglary counts among different census blocks, allowing each block to have its own unique effect on the outcome variable.

Limitations of this model involve the assumption of a linear relationship between the predictors and the log of the expected burglary counts, which may not accurately capture the true relationship in the data. Non-linear relationships or interactions between variables could be overlooked. This model is also assuming that the random effects and errors are normally distributed with constant variances. Violations of these assumptions could affect the validity of the model's inferences.

Results: (reference Table 4 for the coefficients for this model)

The AR(1) structure is applied to the month random effect, with a coefficient ϕ of 0.1525. This indicates a positive correlation between consecutive months, implying that the effect of a month on crime counts is influenced by the previous month.

For fixed effects, year has a moderate negative correlation (-0.199) with the intercept, implying a decreasing trend in burglary counts over time. Moreover, unemployment shows a similar negative correlation (-0.195) with the intercept, suggesting a slight decrease in burglary counts with higher unemployment rates. This is interesting to note as earlier models suggested that higher rates of unemployment suggested higher rates of burglary frequency on Census Blocks.

The minimum and maximum standardized residuals are -1.9445 and 11.3141, respectively. These values indicate the extent to which individual observations deviate from the model's predicted values. The large maximum residual suggests potential outliers or influential observations that may need further investigation.

While the current AR(1) GLMM model exhibits a higher AIC compared to the previous GLMM model (see Table 5), suggesting a potentially weaker fit, both models provide evidence supporting the influence of seasonality on burglary frequency. The higher AIC of the AR(1) GLMM model relative to the other model implies that the latter may be more robust for making inference on the effects of predictors. Nonetheless, the consistent findings across both models underscore the significant impact of seasonal variation on burglary rates

5 Discussion

The methodology used in this paper allowed for the study of the phenomena of crime being susceptible to seasonality. Although the data given was only carried through for five years from 2010 to 2015, we were still able to capture the cyclical effect by treating months as a categorical factor.

The methodology introduced has given the opportunity to address the issues raised in the objectives of the study. Specifically, do characteristics such an unemployment, wealth, or count of young men impact burglary frequency. Given the multicolinearity noted between count of young men and wealth rating, our models were not able to determine the intertwined effect young males have on burglary frequency.

While our analysis recognizes the potential impact of seasonality on burglary frequency,it's important to acknowledge a limitation. Our models fail to answer which months have the most impact on crime. However, through model four, we are able to see through a phi value of 0.1525, there exists a moderate positive autocorrelation between the burglary counts of one month and those of the subsequent month. Moreover,the inclusion of years as a covariate presents a limitation due to the relatively short duration of the study period, encompassing only a five-year span. Consequently, our analysis relies on characteristics specific to Chicago during this time-frame, potentially limiting the generalizability of our findings.

The limitations encountered with our model stem from several factors. Firstly, there exists a potential discrepancy between actual burglary occurrences and reported incidents, as our data relies on individuals reporting crimes. This dependence on reporting behavior could account for the notable frequency of zero counts observed in the burglary distribution (see Figure 1). Additionally, including months as a factor introduces a challenge in interpreting seasonal patterns. The calendar variation inherent in months may generate false indications of seasonal effects, especially since shorter months inherently offer fewer opportunities for criminal activity.

For future research, we recommend acquiring comprehensive datasets that mitigate biases arising from calendar variations. Additionally, expanding the scope of socio-economic characteristics to include ethnicity statistics, healthcare ratings, and levels of police surveillance within Census Blocks would provide richer insights into crime patterns and contribute to more robust analyses.

6 Conclusion

Our study specifically investigates the impact of socio-economic factors on burglary rates within Chicago's Census Block groups. Despite limitations such as the inability to discern specific monthly impacts on crime and the short study duration from 2010 to 2015, our methodology captures that seasonality as a whole makes an impact. Notably, our models highlight the potential influence of seasonality on burglary frequency, although the exact months with the most impact remain unclear. Furthermore, the inclusion of years as a covariate is constrained by the study's limited time-frame, which may restrict the generalizability of our findings to Chicago during this period. These limitations underscore the complexities inherent in analyzing crime data, including discrepancies between reported and actual incidents and challenges associated with interpreting seasonal variations. Nonetheless, our study contributes valuable insights into the relationship between socio-economic characteristics, temporal dynamics, and burglary rates, offering avenues for future research and policy considerations.

7 Appendix

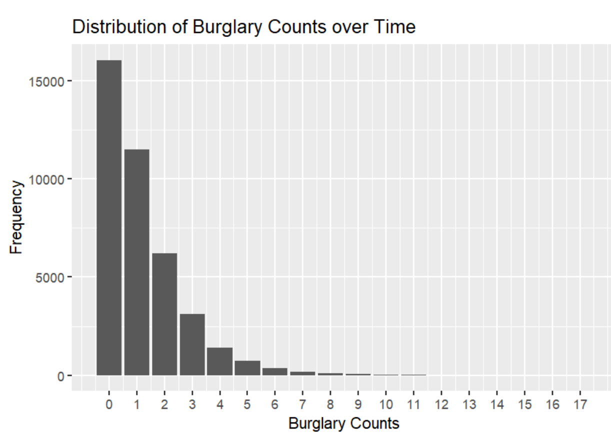

Figure 1: Distribution of Burglary Counts

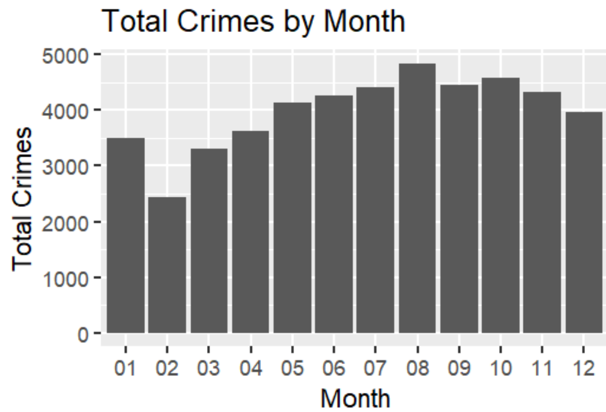

Figure 2: Count of burglaries by month

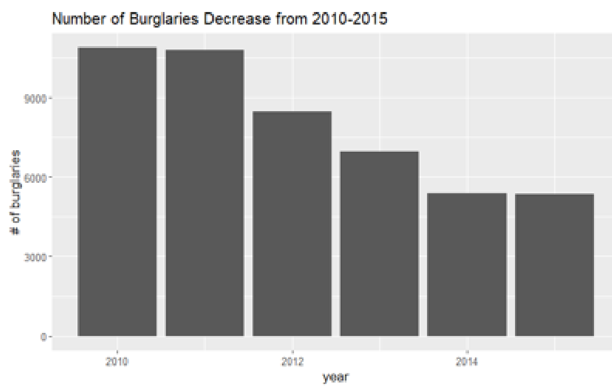

Figure 3: Count of burglaries by month

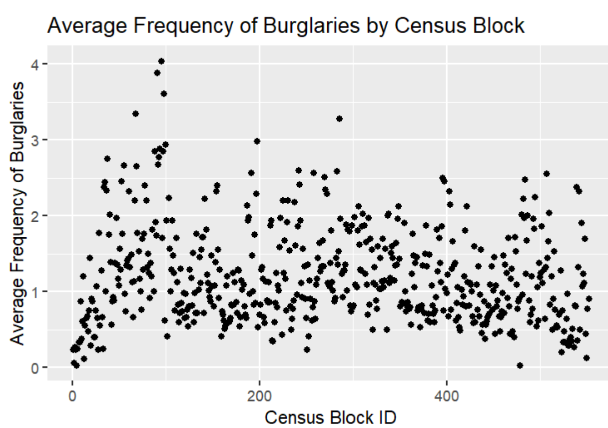

Figure 4: Average Burglaries by Census Block

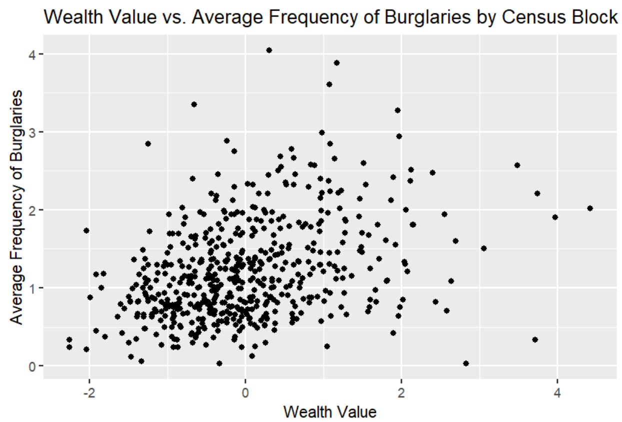

Figure 5: Relationship Between Census Block Wealth Rating and Average Burglary Frequency

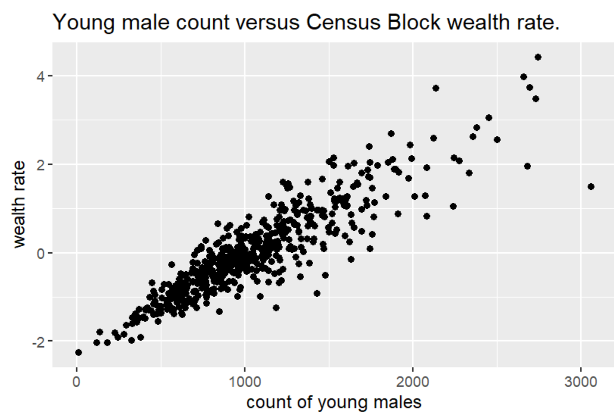

Figure 6: Relationship Between Wealth Rating and Count of Young men for Each Census Block

Predictor	Coefficient	Standard Error	P value
Intercept	331.26	5.52	$< 2 \times 10^{-16}$
wealth_d	0.17	0.00	$< 2 \times 10^{-16}$
year	-0.17	0.00	$< 2 \times 10^{-16}$
un.emp	0.47	0.06	5.54×10^{-14}
month2	-0.36	0.03	$< 2 \times 10^{-16}$
month3	-0.06	0.02	0.018
month4	0.04	0.02	0.133
month5	0.17	0.02	3.67×10^{-13}
month6	0.19	0.02	$< 2 \times 10^{-16}$
month7	0.23	0.02	$< 2 \times 10^{-16}$
month8	0.32	0.02	$< 2 \times 10^{-16}$
month9	0.24	0.02	$< 2 \times 10^{-16}$
month10	0.27	0.02	$< 2 \times 10^{-16}$
month11	0.21	0.02	$< 2 \times 10^{-16}$
month12	0.12	0.02	8.09×10^{-8}

Table 1: Coefficients and P values for Model 1

Predictor	Count Model Coefficient	Zero-Inflation Model Coefficient
(Intercept)	0.470548 ($p < 0.001$)	-1.71973 ($p<0.001$)
wealth_d	1.004043 ($p < 0.001$)	0.64365 ($p < 0.001$)
month2	-0.309791 ($p < 0.001$)	0.23165 ($p = 0.016$)
month3	-0.082350 ($p = 0.006$)	-0.05754 ($p = 0.534$)
month4	0.024704 ($p = 0.387$)	-0.03228 ($p = 0.714$)
month5	0.140106 ($p < 0.001$)	-0.07221 ($p = 0.401$)
month6	0.139016 ($p < 0.001$)	-0.22033 ($p = 0.013$)
month7	0.167931 ($p < 0.001$)	-0.26092 ($p = 0.003$)
month8	0.255741 ($p < 0.001$)	-0.24890 ($p = 0.004$)
month9	0.200265 ($p < 0.001$)	-0.11806 ($p = 0.169$)
month10	0.191587 ($p < 0.001$)	-0.29380 ($p < 0.001$)
month11	0.149990 ($p < 0.001$)	-0.18890 ($p = 0.032$)
month12	0.064435 ($p = 0.021$)	-0.22077 ($p = 0.015$)
year	-0.115776 ($p < 0.001$)	0.19230 ($p < 0.001$)
un.emp	0.797391 ($p < 0.001$)	0.91463 ($p < 0.001$)

Table 2: Count and Zero-inflation model coefficients

Predictor	Estimate	Std. Error	P-value
(Intercept)	0.372036	0.068186	4.86×10^{-8}
wealth_d	0.185827	0.022318	$< 2 \times 10^{-16}$
year	-0.165125	0.002737	$< 2 \times 10^{-16}$
un.emp	0.296227	0.303952	0.33

Table 3: GLMM Coefficients and P-values

Component	Details
Random Effects:	
Intercept	StdDev: 0.1101
Block	StdDev: 0.1098, Corr: 0.647
Month	AR(1) Structure, Phi: 0.1525
Residual	StdDev: 1.4113
Fixed Effects:	
Wealth (wealth_d)	Correlation with Intercept: -0.004
Unemployment (un.emp)	Correlation with Intercept: -0.195
Year	Correlation with Intercept: -0.199
Standardized Residuals:	
Minimum	-1.9445
Maximum	11.3141

Table 4: Important Aspects of AR(1) GLMM Output

Model	AIC score
Model 1- Poisson GLM	120270
Model 2- ZIP	127393
Model 3- GLMM	111635
Model 4- AR1	139324.9

Table 5: Performance of our models with AIC

References

[CF79] Lawrence E Cohen and Marcus Felson. Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4):588–608, 1979.

[CFAP20] Gian Maria Campedelli, Serena Favarin, Alberto Aziani, and Alex Piquero. Disentangling community-level changes in crime trends during the covid-19 pandemic in chicago. *Crime Science*, 9, 10 2020.

[DSD23] R. Delgado and H. Sánchez-Delgado. The effect of seasonality in predicting the level of crime: A spatial perspective. *PLoS One*, 18(5):e0285727, May 2023.

[Fre18] Earl Fredrick. Death, violence, health, and poverty in chicago. *Harvard Public Health Review*, 19:1–25, 2018.

[gov24] Gov. pritzker highlights efforts to reduce violent crime ahead of memorial day weekend. https://www.illinois.gov/news/press-release.24962.html, 2024. Accessed: April 29, 2024.

[LJ10] Michael Lenza and Richard S. Jones. The impacts of political policies, criminality, and money on the criminal justice system in the united states. 2010.

[MLP12] David McDowall, Colin Loftin, and Matthew Pate. Seasonal cycles in crime, and their variability. *Journal of Quantitative Criminology*, 28:389–410, 2012.

[Wei15] David Weisburd. The law of crime concentration and the criminology of place. *Criminology*, 53, 05 2015.