

Rising Assault Rates, Declining Property Crimes, and Uneven Spatial Distributions Across Communities*

Crime Trends in Toronto: A Decade of Analysis (2014–2023)

Weiyang Li

December 23, 2024

This study analyzes a decade of crime data in Toronto (2014–2023) to uncover spatial and temporal patterns of criminal activity across neighborhoods and crime types. By normalizing crime counts per 1,000 residents, we identify a significant rise in assault rates and a decline in property crimes such as break-and-enter incidents. The analysis also reveals substantial spatial disparities, with some neighborhoods experiencing consistently higher per capita crime rates, underscoring localized vulnerabilities. These findings highlight the importance of targeted, data-driven interventions to address rising crime trends and geographic inequities, offering actionable insights for urban policymakers and community stakeholders.

1 Introduction

Urban crime rates are a pressing concern for policymakers, law enforcement agencies, and researchers alike, as they directly impact community safety, economic stability, and quality of life. Studies have long sought to understand crime patterns in urban areas, focusing on factors such as socio-economic disparities, urban infrastructure, and policing strategies. For example, prior research has highlighted the role of concentrated disadvantage in explaining spatial differences in violent crime rates (Sampson, Raudenbush, and Earls 1997) and the temporal patterns of crime influenced by economic downturns (Cantor and Land 1985). Other studies have used machine learning models to predict specific crime events, often with limited interpretability. However, these studies often analyze either spatial or temporal crime patterns in isolation, failing to integrate these two critical dimensions in a unified framework.

*Code and data are available at: <https://github.com/doravmony/static-dynamic-crime>.

This paper addresses these gaps by focusing on crime data from Toronto neighborhoods, leveraging publicly available datasets to study spatial and temporal crime patterns (Toronto 2024). Specifically, the analysis examines variations in crime rates across neighborhoods and crime types, as well as the dynamic changes in these rates over time. While previous research has often focused on high-level trends or predictive modeling, this study incorporates community-specific crime variability and temporal dynamics to provide a more comprehensive understanding. For instance, while some studies have identified broad patterns of urban crime (the “hot spot” theory of crime clustering (Sherman, Gartin, and Buerger 1989)), they often lack specificity in understanding how crime types evolve within communities over time. This research bridges that gap by explicitly modeling crime type, community, and temporal interactions.

Estimand: The primary estimand is the crime rate per 1,000 population for different types of crimes across Toronto neighborhoods and years. This includes estimating community-specific crime rates, the influence of crime type, and temporal trends in these rates.

Despite prior research identifying broad patterns of urban crime, this study fills a critical gap by disentangling the complex interactions between community characteristics, crime types, and temporal trends. Previous studies often focus on one dimension (e.g., spatial or temporal) in isolation or rely heavily on machine learning models that sacrifice interpretability. In contrast, this paper integrates both dimensions using Bayesian hierarchical models, which provide interpretable estimates and capture the variability across neighborhoods and over time. To achieve this, we developed Bayesian hierarchical models, including static and dynamic frameworks, to analyze Toronto’s neighborhood crime data. The static model explores baseline differences across communities and crime types, while the dynamic model incorporates temporal trends and interactions. Additionally, a Random Forest model serves as an alternative approach to validate and contrast results. The posterior predictive checks, convergence diagnostics, and feature importance analyses ensure robust and reliable findings.

Our results reveal significant spatial and categorical differences in crime rates, with some crime types, such as auto theft, being more prevalent in specific communities. Temporal trends show minor but significant year-on-year variations, with certain crime types exhibiting declines. For example, the analysis identifies neighborhoods with persistently high theft rates and demonstrates how these rates have shifted over the past decade. These findings have important implications for public safety policies, enabling targeted interventions tailored to community needs and crime trends.

This paper is structured as follows: Section 2 details the data and preprocessing steps, including feature selection and missing data handling. Section 3 describes the modeling approaches, including the Bayesian hierarchical models and Random Forest. Section 4 presents the results, highlighting key findings from the models and their implications. Section 5 concludes with policy recommendations and broader implications of the study and also discusses the limitations of the analysis and potential avenues for future research.

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023) to analyze crime data collected from the Toronto Open Data Portal (Toronto 2024). The dataset provides detailed yearly crime statistics across Toronto’s neighborhoods from 2014 to 2023, including population counts and reported incidences of various crime categories, such as assault, auto theft, and robbery, spanning multiple years.

The data serves as an important resource for understanding patterns of crime across neighborhoods and over time. In line with previous work on crime analysis (Cantor and Land 1985; Sampson, Raudenbush, and Earls 1997; Sherman, Gartin, and Buerger 1989), this dataset enables us to investigate how regional factors, crime categories, and temporal trends interact to influence crime rates. By normalizing crime counts with population data, we provide a standardized perspective on crime rates per 1,000 people, facilitating meaningful comparisons across regions with differing population sizes.

A preliminary inspection of the dataset reveals variability in crime rates across neighborhoods and over time, highlighting the importance of modeling these patterns using both static and dynamic frameworks. These insights contribute to broader discussions on public safety, regional disparities, and temporal trends in urban crime. Following best practices in statistical storytelling (Alexander 2023), we provide clear and interpretable summaries to communicate our findings effectively to both technical and non-technical audiences.

2.2 Description of Variables

The detailed data cleaning process is provided in the Appendix Section .1, outlining the steps taken to preprocess the raw crime dataset into an analyzable format. After data cleaning, the dataset used for analysis includes the following variables:

Table 1: Description of variables in the cleaned dataset used for analysis, including their names, meanings, and data formats

Table 1: Description of Variables in the Cleaned Dataset

Variable	Description	Format
AREA		
NAME	Name of the neighborhood or area within Toronto.	Character (string)
Year	Year of the reported crime (integer).	Integer
Crime_Type	Type of crime (e.g., ASSAULT, AUTOTHEFT, BIKETHEFT, etc.).	Categorical

Variable	Description	Format
Crime_Count	Total number of reported incidents for a specific crime type in a given year and area.	Numeric
POPULATION		
2023	Population of the area in 2023, used for normalization.	Numeric
Crime_Rate per_1000	Calculated as the number of crimes per 1,000 people, allowing standardized comparisons across areas.	Numeric

The variable **Crime_Rate_per_1000** represents the number of reported crimes of a specific type per 1,000 residents in a given area and year. It was constructed to standardize raw crime counts, which are heavily influenced by population size, enabling fair comparisons across areas. The raw data included crime counts for various crime types (e.g., assault, auto theft, robbery) recorded by year and area, as well as the population for each area in 2023. The data was reshaped from a wide to a long format, creating variables for **Year**, **Crime_Type**, and **Crime_Count**. To account for varying population sizes, crime counts were divided by the population and multiplied by 1,000 to calculate the rate per 1,000 residents. This normalization ensures comparability across areas and facilitates analysis of crime trends over time. The use of crime rates rather than raw counts is particularly important for policy relevance, as it provides a more accurate representation of crime frequency relative to the population, aiding in the development of targeted interventions.

2.3 Alternative Datasets and Rationale for Using the Current Dataset

In exploring potential data sources for this analysis, we considered several alternative datasets, each offering valuable insights but with limitations that made them less suitable for this study’s objectives.

Uniform Crime Reporting (UCR) Dataset

The UCR dataset (Government of Canada, Statistics Canada 2024), provided by Statistics Canada, compiles police-reported crime data across provinces and territories. It includes a wide range of crime categories and is often used for national-level crime trend analyses. However, this dataset aggregates data at provincial and metropolitan levels, which precludes neighborhood-level analysis critical to our research. Furthermore, it lacks key demographic and socioeconomic variables needed to normalize crime rates.

Toronto Police Service (TPS) Data

The Toronto Police Service (Toronto Police Service n.d.) provides public access to raw crime data through its reports and portals. While this dataset overlaps with the one used in this study, it often requires substantial preprocessing to clean and structure the data, as it is not

readily standardized for analysis. Moreover, the TPS dataset does not provide direct links to neighborhood-level population data, which is essential for calculating normalized crime rates.

The dataset selected from the City of Toronto Open Data Portal (Toronto 2024) provides detailed neighborhood-level crime statistics categorized by year and type of crime. This dataset stands out from other potential sources because of its high granularity, consistency, and compatibility with the objectives of this study. Unlike broader datasets such as Statistics Canada’s Uniform Crime Reporting (UCR) dataset, which aggregates crime data at provincial or metropolitan levels, the City of Toronto dataset offers localized information essential for studying intra-city variation in crime rates. Neighborhood-level granularity is a cornerstone of this analysis, as the research aims to understand how crime rates differ across specific communities and evolve over time. The inclusion of crime type breakdowns—such as assault, auto theft, and robbery—enables an exploration of categorical trends, which aligns directly with our goal of analyzing variations between crime categories.

Furthermore, the dataset includes population data for each neighborhood, allowing for the calculation of normalized crime rates per 1,000 residents. This feature is critical for comparability, as it ensures that areas with differing population sizes are evaluated fairly. In contrast, alternative datasets such as the Toronto Police Service (TPS) data lack integrated demographic variables, making it difficult to standardize crime counts across neighborhoods. The TPS data also requires substantial preprocessing to achieve the level of structure already provided by the City of Toronto dataset, which was released in a format ready for analytical workflows.

The dataset’s temporal coverage, spanning multiple years, further supports the study’s objective to investigate dynamic changes in crime rates over time. By contrast, the UCR dataset focuses on aggregated year-over-year trends without providing the fine-grained temporal resolution needed for neighborhood-level analyses. The City of Toronto dataset’s compatibility with temporal and spatial modeling frameworks ensures that we can effectively assess the interaction between community characteristics and temporal effects.

In sum, the City of Toronto dataset is uniquely suited to this research due to its neighborhood-level granularity, categorical breakdowns, temporal consistency, and integrated population data. These features align directly with the study’s focus on spatial and temporal variations in crime rates, enabling rigorous and policy-relevant insights. By leveraging this dataset, the analysis can provide actionable findings that are both methodologically robust and deeply relevant to Toronto’s urban context.

2.4 Measurement

The dataset captures crime as a measurable phenomenon through records of reported incidents, with each entry corresponding to specific types of crimes such as assault, auto theft, and robbery. These data are sourced from police-reported statistics compiled by the City of Toronto and include both temporal and spatial dimensions. Each crime entry is associated with the

year it was reported and the neighborhood where it occurred, enabling a detailed analysis of trends over time and across areas.

From Real-World Events to Dataset Entries

The process begins with an incident being reported to law enforcement, typically by victims, witnesses, or through direct police observation. Once reported, the incident is classified according to established crime categories defined by the Uniform Crime Reporting (UCR) standards, which provide a consistent framework across jurisdictions. For instance, an incident involving forced entry into a residence is categorized as “Break and Enter,” while a violent confrontation resulting in injury is recorded as “Assault.” These standardized classifications ensure that similar incidents are grouped together, facilitating reliable statistical comparisons.

Population data, which is critical for normalizing crime counts into rates, is obtained from municipal records or census estimates. These data allow for the calculation of per-capita crime rates (`Crime_Rate_per_1000`), ensuring meaningful comparisons between areas with differing population sizes. For example, a neighborhood with a higher population might naturally report more incidents; normalizing by population adjusts for this, revealing relative crime risk rather than absolute counts.

Standardization and Limitations

The translation of real-world phenomena into dataset entries involves careful standardization. By adhering to UCR-defined categories and integrating consistent population data, the dataset minimizes potential discrepancies across neighborhoods and years. This standardization ensures that the entries are comparable and meaningful for longitudinal and spatial analyses.

However, this process is not without limitations. Factors such as underreporting, classification inconsistencies, and variability in reporting practices can introduce biases into the dataset. For instance, crimes that go unreported or are misclassified may lead to underestimation or misrepresentation of certain crime categories. Despite these challenges, the dataset provides a robust and systematic framework for analyzing crime trends in Toronto.

Measurement in Context

This approach to measurement reflects the transformation of complex, real-world phenomena into structured data entries suitable for analysis. By integrating both crime and population data, the dataset enables researchers to explore patterns, assess policy impacts, and identify areas of concern within Toronto. While limitations exist, the reliance on standardized data collection and normalization practices ensures that the dataset offers a reliable foundation for studying urban crime dynamics.

2.5 Numerical Summary of Variables

This section focuses on presenting and analyzing the descriptive statistics (e.g., mean, median, standard deviation) and other numerical aspects of the variables.

Table 2: Numerical Summary of Variables

Variable	Mean	Median	Min	Max
Crime Count	35.82	16.00	0	829.0
Crime Rate per 1000	1.87	0.95	0	44.8
Population (2023)	19130.57	18084.00	7057	36388.0

Table 2 presents the numerical summary of the three key variables in the dataset: Crime Count, Crime Rate per 1000, and Population (2023).

- **Crime Count:** The average number of crimes recorded is 35.82, with a median of 16, indicating a right-skewed distribution where most areas report lower crime counts but a few areas have significantly higher counts (up to 829).
- **Crime Rate per 1000:** After normalizing by population, the average crime rate is 1.87 crimes per 1000 people, with a median of 0.95. The crime rate ranges from 0 to 44.8, highlighting variability in per capita crime levels across neighborhoods.
- **Population (2023):** The average neighborhood population is 19,130.57, with a median of 18,084, and the range spans from 7,057 to 36,388. This variability in population underscores the necessity of normalizing crime counts to compare areas meaningfully.

The normalization of crime counts by population helps mitigate the bias introduced by varying population sizes, ensuring a fairer comparison of crime rates across neighborhoods.

2.6 Graphical Summary of Variables

This section highlights visual representations of the data, such as histograms, bar plots, or box plots, to provide insights into the distribution and patterns of the variables.

2.6.1 Outcome variables: Crime Rate per 1000

Figure 1 visualizes the distribution of crime rates per 1,000 residents across neighborhoods in the dataset. The majority of neighborhoods have low crime rates, as evidenced by the high frequency of values concentrated near zero. The distribution is right-skewed, indicating the presence of a small number of neighborhoods with substantially higher crime rates. This skewness highlights the variation in crime rates across different areas, suggesting that while most neighborhoods are relatively safe, a few experience elevated levels of crime. This distribution supports further analysis into factors contributing to these high-crime areas.

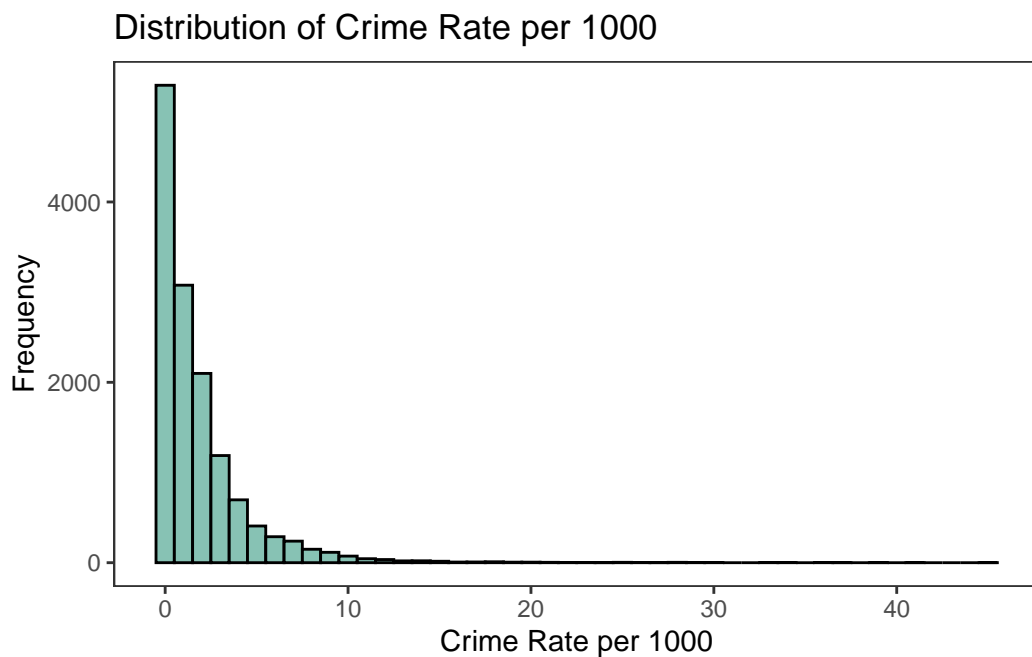


Figure 1: Histogram showing the distribution of crime rates per 1,000 residents across neighborhoods.

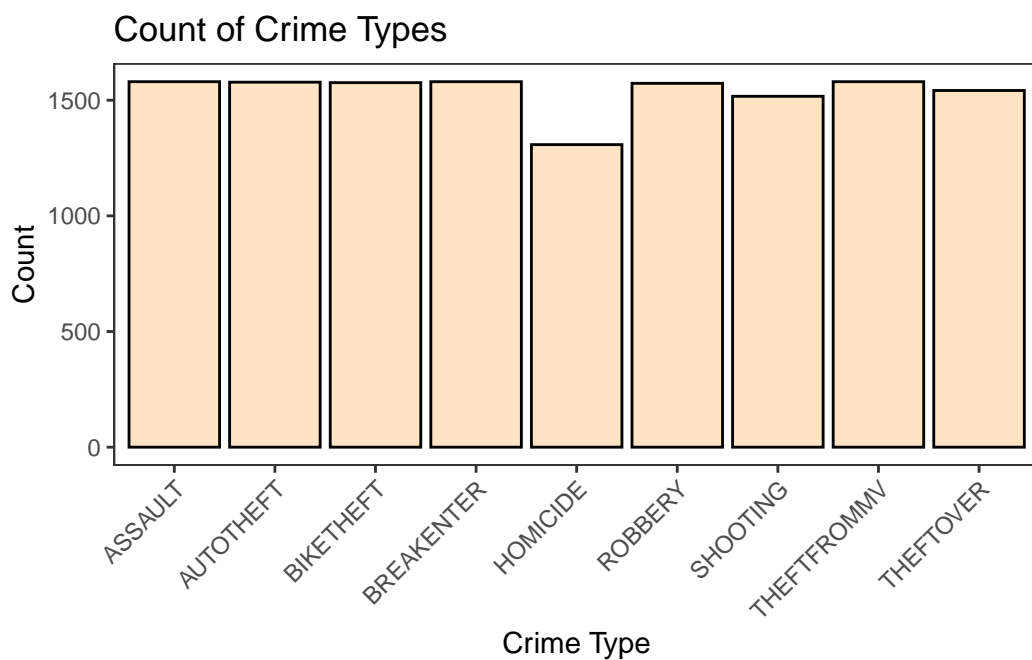


Figure 2: Bar plot displaying the count of recorded instances for each crime type.

2.6.2 Predictor variable: Crime Types

Figure 2 shows the frequency of different crime types recorded in the dataset. The counts are relatively uniform across most crime types, indicating balanced data collection for each category. However, certain crime types, such as **HOMICIDE**, have slightly lower counts compared to others, reflecting their rarity in comparison to crimes like **ASSAULT** or **BREAKENTER**. This uniformity is beneficial for unbiased modeling but underscores the need to account for lower frequencies in rare crime types when interpreting statistical results. The rotated x-axis labels improve readability, especially with long crime type names.

2.6.3 Predictor variable: Population 2023

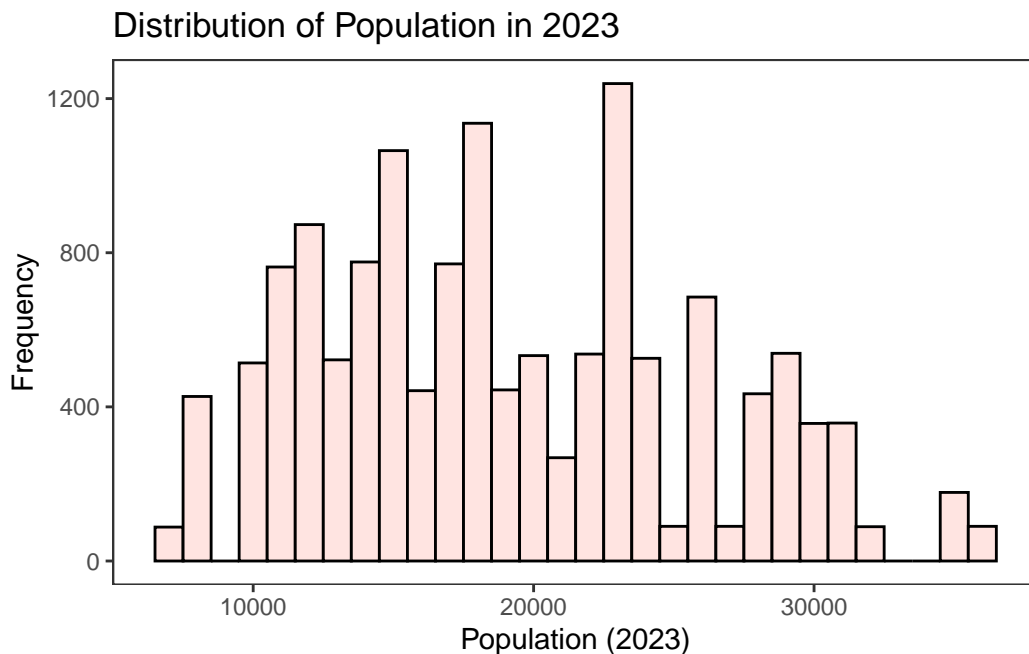


Figure 3: This histogram illustrates the distribution of population sizes in 2023 across different areas, sorted in descending order by frequency.

Figure 3 illustrates the distribution of population sizes in 2023 across different communities. It showcases a relatively uniform spread of population sizes across several intervals, with peaks in the range of approximately 18,000 to 22,000 residents. There is no pronounced right or left skew, indicating a relatively balanced distribution, though extreme values are less frequent. This even distribution suggests a diverse representation of area sizes in the dataset, which is critical for assessing crime rates across varying population scales. The visualization underscores the need to normalize crime counts by population to ensure fair comparisons between areas of different sizes.

2.6.4 Relationship: Crime Count by Year and Crime Type

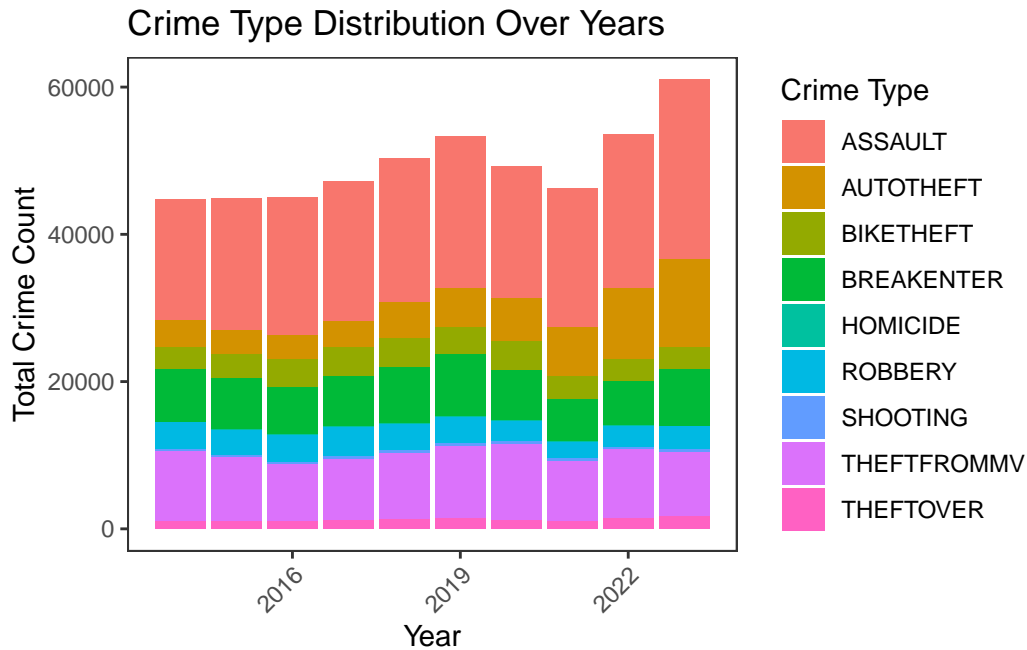


Figure 4: Stacked bar plot showing the yearly distribution of total crime counts across different crime types.

Figure 4 depicts the distribution of total crime counts for each crime type across years. The most prevalent crime type is **ASSAULT**, consistently contributing the largest share across all years. Other crime types, such as **BIKETHEFT** and **BREAKENTER**, also show notable contributions but remain relatively stable compared to **ASSAULT**. Over the years, there is a slight upward trend in total crime counts, indicating an increase in overall criminal activity. The diversity in crime types highlights the multifaceted nature of crime distribution, emphasizing the need for targeted interventions specific to each type.

2.6.5 Relationship: Trends in Crime Rates by Year and Crime Type

Figure 5 illustrates the temporal trends in crime rates per 1000 population across various crime types from 2014 to 2023. **ASSAULT** consistently exhibits the highest average crime rate, with a steady upward trend over the years, highlighting its prevalence and potential escalation. **AUTOTHEFT** shows a notable increase in recent years, surpassing several other crime types. Other crimes such as **BIKETHEFT** and **BREAKENTER** display relatively stable trends, with minor fluctuations. The less frequent crime types, including **HOMICIDE** and **ROBBERY**, exhibit minimal variation over time, suggesting limited temporal changes.

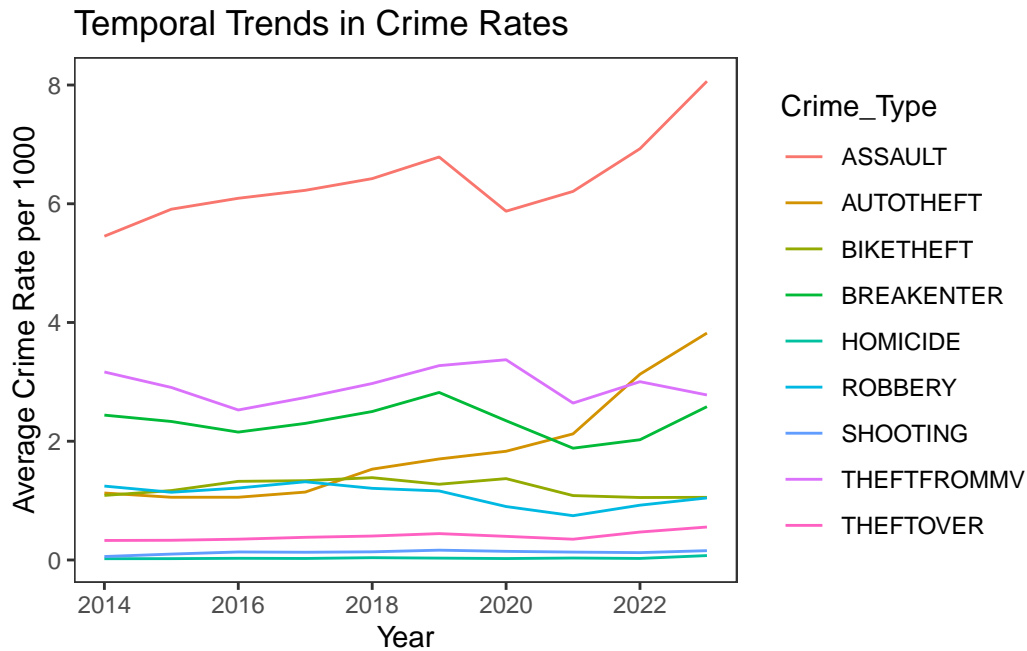


Figure 5: Temporal trends in average crime rates per 1000 population from 2014 to 2023 across different crime types.

2.6.6 Relationship: Crime Rate Trends by Year and Community

Figure 6 displays the temporal trends in average crime rates per 1000 population across different communities from 2014 to 2023. It highlights significant variability between communities. Some communities consistently exhibit higher crime rates, with noticeable upward or downward trends over time, while most remain stable at lower levels. These variations emphasize the importance of considering geographical disparities when analyzing crime patterns. Certain areas, such as those with consistently high crime rates, may warrant further investigation to identify underlying factors or changes in local conditions.

2.6.7 Relationship: Crime Rate Distribution Across Crime Types

Figure 7 visualizes the distribution of crime rates per 1000 population across different crime types. Assaults exhibit the highest median and the widest range, suggesting a significant variation in assault rates between communities. Crimes such as shootings and homicides have relatively low medians and narrower distributions, indicating their rarity but potential concentration in specific areas. Outliers are prevalent across most crime types, reflecting localized spikes in crime rates that deviate from general trends.

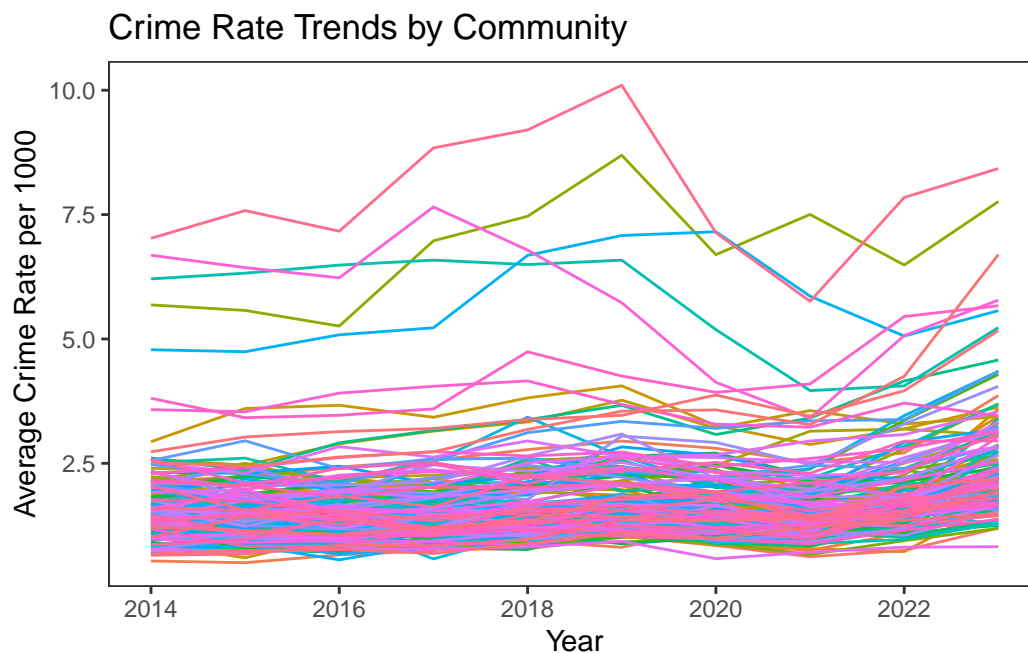


Figure 6: Temporal trends in average crime rates per 1000 population across different communities (2014–2023).

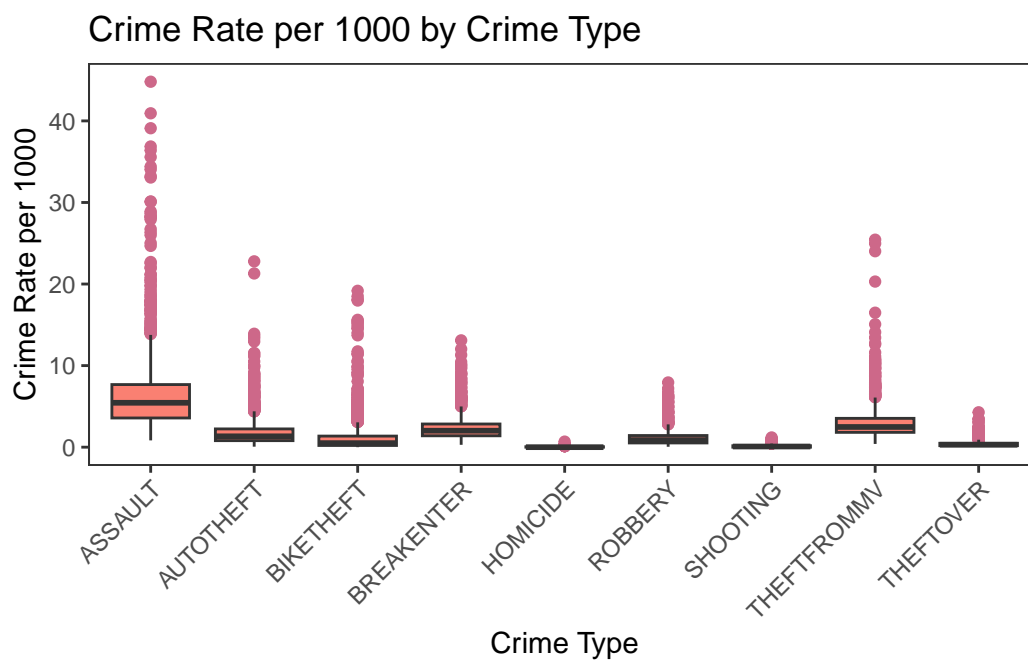


Figure 7: Boxplot showing the distribution of crime rates per 1000 population by crime type.

In the data section, we utilized several R packages to streamline the data processing and visualization workflows. The **tidyverse** package (Wickham et al. 2019) was employed for data manipulation, cleaning, and transformation, including tasks such as selecting relevant columns, reshaping the data into long format, and calculating normalized crime rates per 1000 population. The **dplyr** and **tidyr** packages (Wickham et al. 2023; Wickham, Vaughan, and Girlich 2023), as part of the tidyverse, were instrumental in filtering, pivoting, and mutating data to prepare it for analysis. Additionally, the **arrow** package (Richardson et al. 2024) was used to save the cleaned and processed data efficiently in a Parquet file format, ensuring compatibility with future analyses. Finally, **ggplot2** (Wickham 2016), also part of the tidyverse, was utilized extensively to create visually appealing and informative plots, including histograms, bar plots, and line charts, to effectively communicate insights about crime trends and distributions.

3 Model

To investigate the spatial and temporal patterns of crime rates across communities, we employ a Bayesian hierarchical mixed-effect modeling approach that integrates both static and dynamic components. Background details and diagnostics of the models are included in Appendix-Section [.2](#).

3.1 Model set-up

To address the research questions, we construct Bayesian hierarchical mixed-effect models for static and dynamic components. Both models use `Crime_Rate_per_1000` as the response variable (normalized crime rate), allowing for meaningful comparisons between communities and across years. The models are implemented using weakly informative priors and assume Gaussian distribution for the response variable, reflecting its continuous nature.

3.1.1 Complete Model Formula

3.1.1.1 Response Variable

$$Y_{i,j,t} = \text{Crime Rate per 1000 for community } i, \text{ crime type } j, \text{ year } t$$

3.1.1.2 Static Model

The static model focuses on identifying significant differences in crime rates across communities and crime types:

$$Y_{i,j} = \text{Gaussian}(\mu_{i,j}, \sigma^2)$$

$$\mu_{i,j} = \beta_0 + \gamma_j + b_i$$

Where:

- β_0 : Global mean crime rate.
- γ_j : Fixed effect for crime type j (e.g., theft or assault).
- b_i : Random effect for community i, representing deviations from the global mean ($b_i \sim N(0, \sigma_b^2)$).

3.1.1.3 Dynamic Model

The dynamic model incorporates temporal changes and community-specific trends:

$$Y_{i,j,t} = \text{Gaussian}(\mu_{i,j,t}, \sigma^2)$$

$$\mu_{i,j,t} = \beta_0 + \gamma_j + \delta_t + (\gamma_j \cdot \delta_t) + b_i + b_{i,t}$$

Where:

- β_0 : Global mean crime rate.
- γ_j : Fixed effect for crime type j.
- δ_t : Fixed effect for year t, representing overall temporal trends.
- $\gamma_j \cdot \delta_t$: Interaction effect between crime type and year.
- b_i : Community-specific random effect ($b_i \sim N(0, \sigma_b^2)$).
- $b_{i,t}$: Community-year interaction (random slope for year; $b_{i,t} \sim N(0, \sigma_{b_t}^2)$).

3.1.1.4 Priors

We use weakly informative priors for all parameters to balance interpretability and flexibility:

- **Fixed effect:** $\beta_0, \gamma_j, \delta_t \sim N(0, 10)$. These priors assume the effects are centered around zero, with a wide range to account for potentially large or small effects without being overly restrictive.
- **Random effects:** Community-specific effects $b_i \sim N(0, \sigma_b^2)$ and Community-year interactions $b_{i,t} \sim N(0, \sigma_{b_t}^2)$. These priors reflect the expectation that deviations from the global mean are normally distributed, with variances estimated from the data.

- **Residual variance:** $\sigma^2 \sim \text{Cauchy}(0, 2)$. The Cauchy prior accommodates greater variability in residuals, providing robustness to outliers or high noise levels.

Detailed prior distributions, their visualizations, and sensitivity analyses are provided in Appendix-[Section .2.1].

We run the model in R (R Core Team 2023) using the **brms** package of Bürkner (2021).

3.1.2 Model justification

We expect community-specific factors and temporal trends to influence crime rates across different crime types. In particular, the hierarchical structure of the data, with observations nested within communities and years, motivates the use of a Bayesian hierarchical mixed-effect model.

Specifically:

- **Crime type:** Different crime types exhibit distinct baseline rates. Including crime type as a fixed effect allows the model to capture systematic differences in crime rates between these categories.
- **Temporal trends:** Crime rates are expected to vary over time, potentially increasing or decreasing depending on broader societal and economic trends. Adding year as a fixed effect captures these overall temporal changes, while community-specific random slopes account for localized deviations in trends.
- **Community effects:** Communities may have unique characteristics that affect baseline crime rates. Treating communities as random effects helps account for this unobserved heterogeneity.
- **Interaction between crime type and year:** Certain crimes might increase or decrease more rapidly than others over time. Including an interaction term allows the model to capture these varying temporal dynamics across crime types.

3.1.3 Modeling Decisions and Data Characteristics

Our modeling approach reflects careful consideration of the characteristics and structure of the data, as outlined in the data section. Key decisions about feature inclusion, transformations, and hierarchical structures were made to ensure the model aligns with the data’s unique properties and research objectives.

- **Use of Standardized Crime Rates:** Crime counts vary significantly across communities due to differences in population size. Directly comparing crime counts without accounting for population differences would misrepresent relative crime risks. We use `Crime_Rate_per_1000`, a standardized crime rate (crimes per 1,000 residents), as the response variable instead of `Crime_Count`. This ensures that differences in crime levels reflect actual risk rather than population size, making comparisons across communities meaningful.
- **Inclusion of Crime Type as a Fixed Effect:** The dataset categorizes crimes into distinct types (e.g., theft, assault), and each crime type may have unique dynamics and spatial distributions. We include `Crime_Type` as a fixed effect to capture systematic differences in crime rates by type. This allows the model to identify whether certain types of crimes are more prevalent in specific communities or show distinct temporal trends.
- **Incorporation of Year and Temporal Trends:** The dataset spans multiple years, providing an opportunity to analyze temporal trends in crime rates. These trends may vary by community or crime type. In the **Dynamic Model**, we include `Year` as both a fixed and random effect:
 - **Fixed Effect:** Captures overall temporal trends, such as whether crime rates are increasing or decreasing over time.
 - **Random Slope for Year:** Allows temporal trends to vary across communities, accounting for region-specific patterns of change.
- **Treatment of Communities as Random Effects:** The dataset contains multiple communities, each with unique characteristics that influence crime rates. Modeling these differences explicitly is critical to understanding spatial variation in crime patterns. Communities are treated as random effects:
 - In the **Static Model**, we include $(1|\text{AREA_NAME})$ to capture community-specific baseline differences in crime rates.
 - In the **Dynamic Model**, we extend this to $(1|\text{YEAR} + \text{AREA_NAME})$, allowing community-specific intercepts and slopes to capture both baseline differences and temporal trends.
- **Gaussian Assumption for Crime Rates:** The response variable, `Crime_Rate_per_1000`, is continuous and normalized. While crime counts often follow a Poisson distribution, the transformation into a continuous rate justifies using a Gaussian distribution for the response. We assume a Gaussian distribution for `Crime_Rate_per_1000`. This simplifies model interpretation while still providing accurate predictions for continuous, standardized data.

3.1.4 Assumptions, Limitations, and Applicability

Our model is based on several underlying assumptions about the data and its structure. First, we assume that the response variable, `Crime_Rate_per_1000`, follows a Gaussian distribution. This assumption is reasonable for normalized, continuous crime rates but may not fully capture extreme variations or skewness in the data. Additionally, the model assumes additive relationships between predictors, such as crime type and year, meaning that their effects combine linearly. Non-linear relationships or interactions beyond those explicitly included in the model are not captured.

The model also assumes that the random effects, such as community-specific effects and community-year interactions, are independent and normally distributed. While this assumption simplifies the hierarchical structure, it may overlook potential correlations between communities, especially geographic or socio-economic dependencies. Finally, we use weakly informative priors, such as $N(0, 10)$ for fixed effects, which reflect limited prior knowledge while allowing the data to dominate the posterior estimation.

Despite these strengths, the model has several limitations. Temporal effects are modeled linearly, which may not adequately represent crime rates that exhibit sudden spikes, drops, or other non-linear trends. Similarly, the model does not explicitly account for spatial correlations between communities, which could bias estimates if nearby communities share similar trends. Furthermore, the current model abstracts away the count-specific nature of crime data by focusing on normalized rates, potentially missing important characteristics like over-dispersion or zero inflation.

This model may not be appropriate in certain scenarios. For instance, if crime rates exhibit strong skewness or heavy tails, a Gaussian likelihood may not be suitable, and alternative distributions like log-normal or gamma may be required. Similarly, for data with significant zero inflation or overdispersion (e.g., rare crimes in small communities), a Poisson or Negative Binomial model would be more effective. Lastly, in cases where spatial dependencies or non-linear temporal trends are significant, extensions such as spatial models or non-linear temporal terms would provide better results.

In summary, while the model is well-suited for analyzing hierarchical, temporal, and community-specific patterns in crime rates, it may require adjustments or extensions to address specific data characteristics, such as spatial correlations, non-linear trends, or extreme distributions. These considerations ensure that the model remains robust and interpretable across diverse datasets and research contexts.

3.1.5 Software and Model Validation

The models were implemented using the `brms` package (Bürkner 2021) in R, which interfaces with the Bayesian modeling framework Stan. This setup allowed us to specify and fit hierarchical mixed-effect models efficiently using Hamiltonian Monte Carlo sampling. The use of `brms`

provided robust tools for defining priors, estimating posterior distributions, and performing diagnostic checks, ensuring both flexibility and reliability in our modeling approach.

To validate the models, we employed a range of techniques to assess predictive performance and robustness. The dataset was split into training (70%) and testing (30%) subsets. Models were fitted using the training data, and predictions were made for the test set to evaluate out-of-sample performance. Root Mean Square Error was calculated as the primary metric, enabling a quantitative comparison of predicted and observed crime rates. Both the static and dynamic models demonstrated competitive RMSE values, indicating strong predictive capabilities.

Posterior predictive checks were conducted to assess how well the models replicated the observed data. These checks confirmed that the models captured the main patterns and trends in crime rates while avoiding overfitting. Additionally, sensitivity analyses were performed by varying prior distributions and re-estimating the models. The results remained consistent across different prior specifications, demonstrating the robustness of the posterior estimates and the stability of the modeling framework.

Convergence diagnostics further supported the validity of the models. Standard metrics, including \hat{R} statistics (all $\hat{R} \approx 1$), confirmed proper convergence across all parameters. Visual inspection of trace plots showed no evidence of sampling issues, indicating that the posterior distributions were reliably estimated. Full details, including diagnostic plots, RMSE calculations, and additional validation results, are included in the Appendix to provide comprehensive evidence of the models' reliability and performance.

3.2 Alternative Model

As an alternative to the Bayesian hierarchical mixed-effect models, we considered a Random Forest model to predict normalized crime rates. Random Forest is a machine learning approach that excels at capturing complex, non-linear relationships and interactions between predictors without requiring explicit model specification. It operates by aggregating predictions from an ensemble of decision trees, reducing the risk of overfitting.

3.2.0.1 Strengths of Random Forest

- **Non-linear Modeling:** Unlike the Bayesian models, Random Forest can naturally handle non-linear relationships and complex interactions between variables, such as interactions between community characteristics, crime types, and temporal trends.
- **Flexibility:** The method does not rely on distributional assumptions (e.g., Gaussian likelihood) for the response variable, making it robust to outliers or skewed data.
- **Variable Importance:** Random Forest provides measures of variable importance, helping identify the predictors that contribute the most to the model's performance.

3.2.0.2 Weaknesses of Random Forest

- **Interpretability:** While Random Forest offers strong predictive performance, its “black-box” nature makes it difficult to interpret relationships between predictors and outcomes. Unlike Bayesian models, it does not provide explicit estimates of fixed and random effects or quantify uncertainty in predictions.
- **Lack of Hierarchical Structure:** Random Forest does not explicitly account for the nested structure of the data (e.g., years within communities), which could lead to sub-optimal performance in hierarchical datasets.
- **Overfitting Risks:** Although Random Forest reduces overfitting through tree aggregation, it may still struggle with small sample sizes or imbalanced datasets.

3.2.0.3 Performance Comparison

We trained the Random Forest model on the same training dataset used for the Bayesian models and evaluated its performance on the test set. While Random Forest achieved a slightly lower RMSE compared to the Bayesian models, it lacked the ability to explicitly model community-specific random effects, temporal trends, and their interactions.

3.2.0.4 Rationale for Final Model Choice

The Bayesian hierarchical mixed-effect models were chosen as the primary modeling framework due to their interpretability, flexibility in representing the hierarchical structure of the data, and ability to quantify uncertainty. These models allow us to directly estimate the effects of crime type, temporal trends, and community-specific variations, which align with the research objectives of understanding spatial and temporal crime patterns. While Random Forest provided a useful alternative for evaluating non-linear relationships and variable importance, its lack of interpretability and hierarchical structure made it less suitable for addressing the specific goals of this study.

4 Results

4.1 Static Model

Table 3: Summary of static model results with fixed effects for crime type and 95% credible intervals.

Table 3: Static Model Results

Parameter	Estimate	Est.Error	1.95..CI	u.95..CI
Intercept	6.39	0.10	6.19	6.57

Parameter	Estimate	Est.Error	1.95..CI	u.95..CI
Crime_TypeAUTOTHEFT	-4.53	0.07	-4.68	-4.38
Crime_TypeBIKETHEFT	-5.20	0.07	-5.35	-5.06
Crime_TypeBREAKENTER	-4.06	0.08	-4.21	-3.91
Crime_TypeHOMICIDE	-6.41	0.08	-6.56	-6.26
Crime_TypeROBBERY	-5.27	0.08	-5.42	-5.12
Crime_TypeSHOOTING	-6.25	0.08	-6.39	-6.09
Crime_TypeTHEFTFROMMV	-3.43	0.07	-3.58	-3.29
Crime_TypeTHEFTOVER	-6.00	0.07	-6.14	-5.86
sigma	1.75	0.01	1.73	1.78

Table 3 provides a summary of the static model’s estimates and diagnostic metrics. This model examines baseline differences in crime rates across crime types while accounting for community-specific variability.

- The **Intercept (6.39)** represents the average baseline crime rate per 1,000 population for the reference crime type in the reference community.
- The fixed effects for crime types are all negative, indicating that the crime rates for these types are lower compared to the reference crime type. For example:
 - **AUTOTHEFT**: The estimated coefficient (-4.53) suggests that the crime rate for auto theft is 4.53 fewer crimes per 1,000 population than the reference crime type.
 - **HOMICIDE**: The estimate (-6.41) indicates that homicides occur at a significantly lower rate than the reference crime type.
- The **random intercept (sd(Intercept)=1.05)** shows that some communities deviate considerably from the overall baseline crime rates, highlighting the importance of community-level factors.
- The residual standard deviation ($\sigma = 1.75$) reflects the unexplained variability, suggesting that while crime type and community effects are important, other factors (e.g., socio-economic variables, policing) likely influence crime rates.

The static model estimates crime rates using fixed effects for crime types and a random intercept for communities. Fixed effects reveal significant differences between crime categories, with coefficients representing deviations from the reference crime type. The random intercept indicates substantial variability between communities, even after accounting for crime types, while the residual standard deviation suggests unmodeled factors contribute to the remaining variability. Diagnostics show good convergence and sufficient effective sample sizes, supporting the model’s robustness in explaining spatial and categorical differences.

4.2 Dynamic Model

Table 4: Summary of dynamic model results with fixed effects, interactions, and 95% credible intervals.

Table 4: Dynamic Model Results

Parameter	Estimate	Est.Error	l.95..CI	u.95..CI
Intercept	-99.51	12.60	-123.87	-74.43
Crime_TypeAUTOTHEFT	-32.27	9.97	-51.42	-12.87
Crime_TypeBIKETHEFT	7.89	9.83	-11.58	27.30
Crime_TypeBREAKENTER	8.91	9.75	-9.91	27.76
Crime_TypeHOMICIDE	6.06	10.04	-14.06	25.46
Crime_TypeROBBERY	11.84	9.65	-6.72	30.06
Crime_TypeSHOOTING	5.22	9.92	-13.71	24.52
Crime_TypeTHEFTFROMMV	6.79	9.58	-11.69	25.68
Crime_TypeTHEFTOVER	4.15	9.64	-14.70	23.25
Year	0.05	0.01	0.04	0.06
Crime_TypeAUTOTHEFT:Year	0.01	0.00	0.00	0.02
Crime_TypeBIKETHEFT:Year	-0.01	0.00	-0.02	0.00
Crime_TypeBREAKENTER:Year	-0.01	0.00	-0.02	0.00
Crime_TypeHOMICIDE:Year	-0.01	0.00	-0.02	0.00
Crime_TypeROBBERY:Year	-0.01	0.00	-0.02	0.00
Crime_TypeSHOOTING:Year	-0.01	0.00	-0.02	0.00
Crime_TypeTHEFTFROMMV:Year	-0.01	0.00	-0.01	0.00
Crime_TypeTHEFTOVER:Year	-0.01	0.00	-0.01	0.00
sigma	1.74	0.01	1.72	1.77

Table 4 provides a summary of the dynamic model’s estimates and diagnostic metrics. This model extends the static approach by incorporating temporal trends and interactions between crime types and year, while also accounting for community-level effects over time.

- The **Intercept (-99.51)** and Year (0.05) coefficients indicate an upward trend in crime rates over time, starting from a very low baseline. The intercept is negative because the model includes year as a numeric variable, and the starting point (e.g., Year = 2014) leads to a low baseline prediction.
- The interaction terms between crime type and year reveal subtle temporal patterns:
 - **AUTOTHEFT**: The positive interaction (0.01) suggests a slight increase in the rate of auto theft over time.

- **BIKETHEFT**: The negative interaction (-0.01) indicates a small decline in bike theft rates over time.
- The **random intercept** ($\text{sd}(\text{Intercept})=0.97$) shows that community-specific baseline differences persist, albeit with slightly less variability than in the static model.
- The **random slope for year** ($\text{sd}(\text{Year})=0.00$) suggests that temporal trends are consistent across communities, with negligible variation.

The dynamic model extends the static model by incorporating temporal effects (Year) to capture overall trends over time, interactions between crime type and year to assess how trends vary across crime categories, and a random slope for year within communities to account for community-specific temporal variations. The fixed effect for year indicates a small but statistically significant increase in baseline crime rates over time, while the interaction terms between crime type and year are mostly small and negative, suggesting slight declines in crime rates for most crime types. The random slope for year shows that temporal trends are nearly identical across communities, with minimal variation. Model diagnostics confirm good convergence for most parameters, though the random slope parameters exhibit slightly higher uncertainty due to small effective sample sizes.

4.3 Alternative Model

Table 5: Feature importance from the Random Forest model based on Increase in Node Purity.

Table 5: Feature importance from the Random Forest model

Feature	Increase in Node Purity
Crime_Type	30626.03
Year	1276.36
AREA_NAME	6155.32

Table 5 shows the feature importance results from the Random Forest model, measured by the Increase in Node Purity, provide insights into how much each feature contributes to reducing prediction error. Here’s an analysis of the results:

- The feature Crime_Type has the highest importance (30626.03), indicating that the type of crime is the most critical factor in predicting crime rates. This is expected because crime types inherently capture significant differences in baseline rates, trends, and spatial patterns. The overwhelming importance of **Crime_Type** aligns with findings from the static and dynamic Bayesian models, where crime types significantly influence crime rates. This underscores that the type of crime drives most of the variability in the data.

- The **AREA_NAME** variable, representing the community, has a moderate importance score (6155.32). This reflects that community-specific factors, such as socio-economic conditions or policing strategies, play a significant role in explaining crime rate variability across regions. The importance of **AREA_NAME** highlights the spatial variability in crime rates, consistent with the random intercepts used in the Bayesian models. It shows that local factors specific to each community are important predictors of crime.
- The variable **Year** has the lowest importance score (1276.36) compared to the other features. This suggests that temporal trends contribute relatively less to predicting crime rates, indicating that crime rates may be relatively stable over time, with fewer year-on-year variations. The relatively low importance of **Year** supports the idea that temporal trends have a limited impact compared to spatial and categorical factors. This finding is consistent with the dynamic Bayesian model, where the fixed effect for **Year** was small but statistically significant.

4.4 Comparison

Table 6: Comparison of Model Performance Based on RMSE.

Table 6: Comparison of Model Performance

Model	RMSE
Random Forest	1.876700
Static Model	1.745739
Dynamic Model	1.738076

Table 6 shows the comparison between models based on the root mean squared error:

- The dynamic model achieves the lowest RMSE (1.738), indicating the best overall predictive accuracy. This is expected as the dynamic model accounts for temporal trends and interactions, providing a richer structure for capturing variability in crime rates.
- The static model has an RMSE of 1.746, which is very close to the dynamic model. This suggests that much of the variability is explained by spatial and categorical factors rather than temporal trends.
- The Random Forest model has the highest RMSE (1.877), indicating slightly lower predictive performance compared to the Bayesian models. While Random Forest is flexible and effective for capturing non-linearities, it may struggle with hierarchical data structures and interpretability.

4.5 Posterior predictive check

Posterior predictive checks (PPCs) were performed to evaluate the goodness-of-fit for both the static and dynamic models, assessing how well the predicted distributions align with the observed data.

4.5.1 Static Model

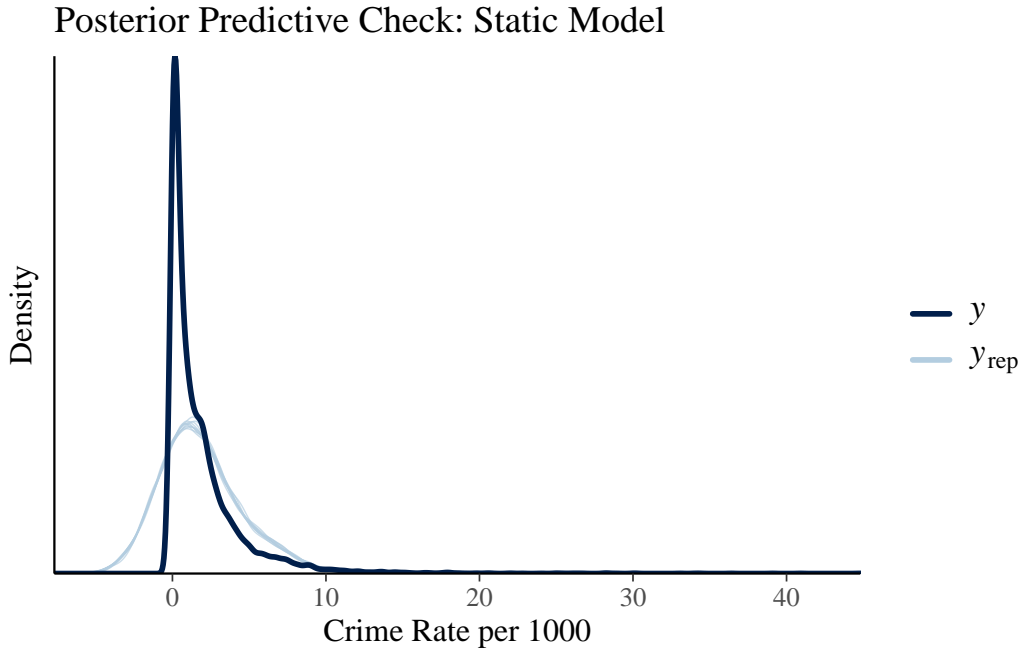


Figure 8: Posterior predictive check for the static model: Observed and simulated densities of crime rates, showing the model's ability to capture the overall data distribution.

This posterior predictive check as shown in Figure 8 for the static model compares the observed crime rate density (y) with the model's predicted density (y_{rep}). The observed density (dark line) exhibits a sharper peak and longer tail compared to the predicted density (light line). This indicates that while the model captures the general trend of crime rate distribution, it may underestimate extreme values or variability present in the observed data. Adjustments in the model specification or inclusion of additional explanatory variables could improve alignment with observed patterns.

The posterior predictive check histogram as shown in Figure 9 for the static model compares the observed (y , dark bars) and predicted (y_{rep} , light bars) distributions across multiple replicated datasets. The observed distribution shows a sharper peak and heavier tail compared to the predicted distributions, indicating that the static model struggles to fully capture the

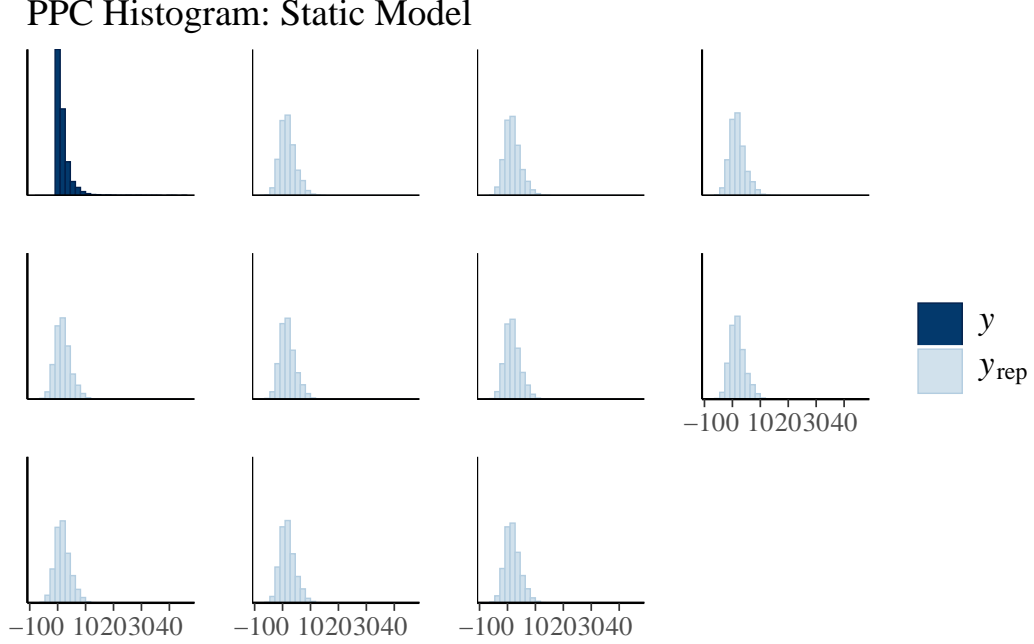


Figure 9: Posterior predictive check for the static model: Comparison of observed and simulated crime rates using a histogram.

variability and extreme values in the observed data. This highlights potential model limitations in addressing the underlying heterogeneity in crime rates.

4.5.2 Dynamic Model

The posterior predictive check as shown in Figure 10 for the dynamic model compares the observed crime rate density (y) with the predicted density (y_{rep}). Similar to the static model, the observed density (dark line) is more sharply peaked and exhibits a longer tail than the predicted density (light line). However, the dynamic model slightly improves the alignment of the central tendency of the observed data. The dynamic model captures temporal trends and interactions but still underestimates the extreme variability in the observed crime rates, indicating potential areas for further model refinement.

Figure 11 compares observed crime rates (y) against the average predicted crime rates (y_{rep}) from the dynamic model. The dashed line represents perfect alignment, where predictions would match observations exactly. While most points cluster near the line for lower crime rates, there is greater dispersion as crime rates increase, suggesting that the model performs well for smaller values but struggles with higher crime rate predictions. This indicates a potential area for improvement in capturing extreme variations.

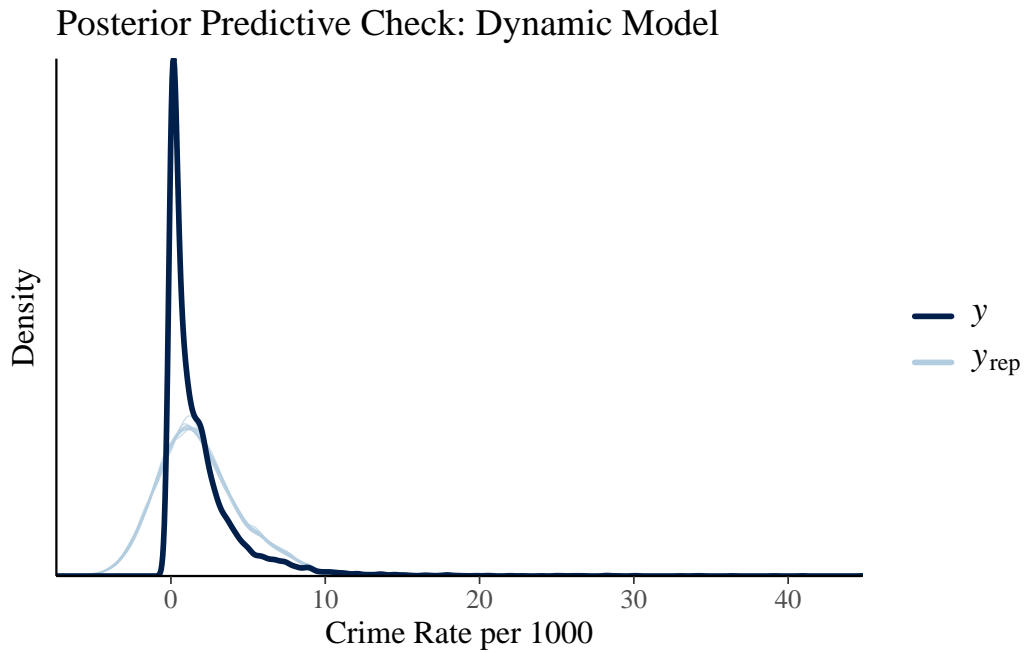


Figure 10: Posterior predictive check for the dynamic model: Observed and simulated densities of crime rates, highlighting the model's fit with temporal and community-level variability.

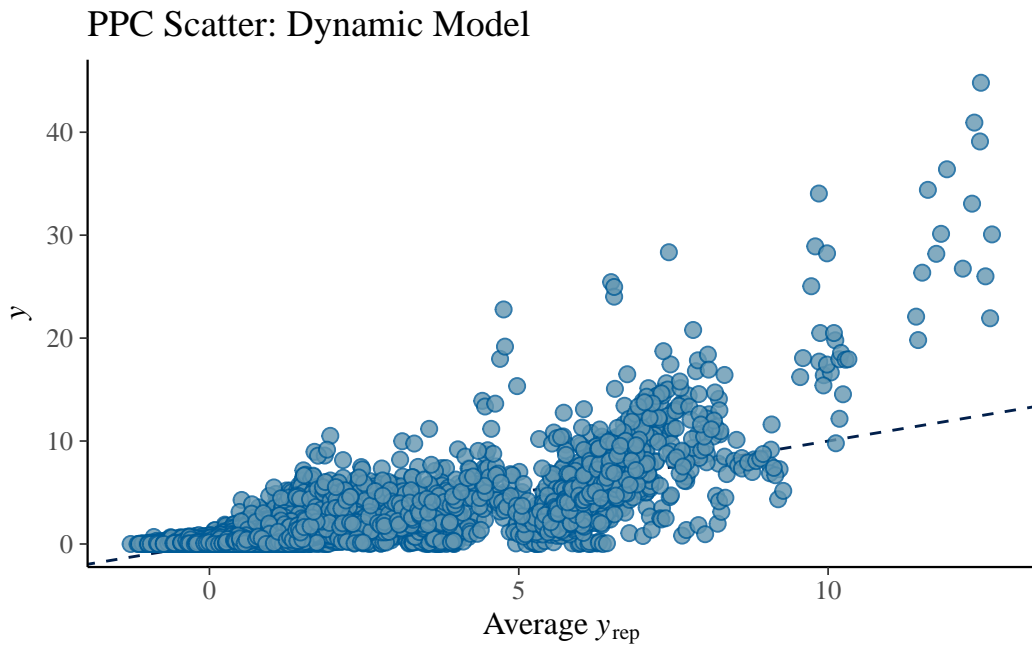


Figure 11: Posterior predictive check for the dynamic model: Scatterplot showing the relationship between observed and predicted average crime rates.

We used Bayesian hierarchical models for analysis, assigning weakly informative priors to all parameters. The priors ensure sufficient regularization without overly constraining the model. Full details, including prior visualizations and sensitivity checks, can be found in Appendix-[Section .2.1]. Posterior predictive checks above demonstrate that both models replicate the observed data distributions well. Comprehensive diagnostics, including convergence checks and sensitivity analyses, are provided in Appendix-[Section .2.2].

In the results section, we used the `brms` package (Bürkner 2021) in R to implement Bayesian hierarchical models, perform posterior predictive checks, and evaluate model fit using diagnostic metrics. The `caret` package (Kuhn and Max 2008) was used to split the data into training and testing sets, ensuring a robust evaluation of model performance. Additionally, a Random Forest model was developed using the `randomForest` package (Liaw and Wiener 2002) to provide an alternative non-parametric comparison.

5 Discussion

5.1 Overview of This Study

This study provides an in-depth examination of crime trends in Toronto over nearly a decade, leveraging police-reported data and population statistics. The analysis focuses on understanding the temporal evolution of crime rates, spatial disparities across neighborhoods, and variations by crime type. By normalizing crime counts into rates per 1,000 people, the study accounts for population differences, offering a more accurate and comparable measure of crime levels. Through data cleaning, visualization, and statistical modeling, this research uncovers critical patterns that inform our understanding of urban crime dynamics.

5.2 Key Insights on Temporal and Spatial Crime Trends

One of the key findings of this study is the temporal variation in crime rates. Certain crime types, such as assault and auto theft, have shown consistent increases over the years, potentially reflecting broader societal shifts, including economic instability, urbanization, or changing policing strategies. Conversely, crimes such as break-and-enter have seen a steady decline, which may be attributed to advancements in home security technologies, increased surveillance, or targeted enforcement efforts.

Spatial analysis reveals substantial disparities in crime rates across Toronto's neighborhoods. Areas with higher population densities or socio-economic vulnerabilities often exhibit elevated crime rates, while more affluent or less accessible neighborhoods tend to report lower rates. These spatial variations highlight the importance of addressing localized factors such as poverty, unemployment, and access to public services in crime prevention efforts.

5.3 Crime Type-Specific Observations

Breaking down the data by crime type offers valuable insights into the underlying factors influencing different forms of crime. Property crimes such as theft and auto theft are often tied to economic conditions, suggesting that financial hardships can drive individuals toward opportunistic crimes. Violent crimes like assault and robbery, on the other hand, may be more closely linked to social tensions, personal conflicts, or alcohol and substance abuse.

By normalizing crime data, the study also reveals that smaller neighborhoods can experience disproportionately high crime rates per capita, even if their absolute crime counts are low. This finding underscores the vulnerability of such areas, where fewer resources may be available for law enforcement and community support. Recognizing these vulnerabilities is essential for designing equitable and effective crime reduction programs.

5.4 Limitations of the Study

This study's strength lies in its use of normalized crime rates, which provides a fairer basis for comparison across neighborhoods of varying populations. The use of a longitudinal dataset spanning multiple years allows for the identification of meaningful trends, while the incorporation of spatial dimensions enhances the understanding of crime distribution within the city.

However, the study is not without limitations. The reliance on police-reported data introduces potential biases, such as underreporting, particularly for crimes like sexual assault or minor theft, which are often less likely to be reported. Additionally, variations in how crimes are classified or recorded across different years or jurisdictions can impact data consistency.

Another limitation is the absence of contextual variables, such as socio-economic indicators, unemployment rates, or education levels, which could provide deeper insights into the causes of observed trends. The lack of qualitative data, such as interviews with law enforcement or community members, further limits the interpretation of findings and understanding of the lived experiences of crime-affected populations.

5.5 Future Directions and Areas for Improvement

Future research should strive to address these limitations by integrating additional datasets, such as economic, social, and demographic indicators, to build a more comprehensive picture of crime dynamics. Incorporating qualitative methods, such as surveys or focus groups, could provide contextual insights into the drivers of crime and the effectiveness of current prevention strategies.

Advanced statistical and computational techniques, such as machine learning models or spatial econometrics, could help uncover hidden patterns and improve the predictive accuracy of crime

trend analyses. Exploring cross-city or cross-regional comparisons could also shed light on whether observed trends are unique to Toronto or reflective of broader urban phenomena.

Moreover, future studies could focus on specific interventions, such as evaluating the impact of increased police presence or community engagement initiatives on crime rates. Identifying the most effective measures could help allocate resources more efficiently and equitably.

5.6 Conclusion

This study sheds light on the complex dynamics of crime in Toronto, offering valuable insights into temporal trends, spatial disparities, and the role of different crime types. By normalizing crime rates and focusing on population-adjusted measures, the research provides a robust foundation for understanding urban crime. However, addressing data limitations and expanding the scope of analysis will be critical for advancing crime prevention strategies and fostering safer communities in the future.

Appendix

.1 Data cleaning

Step 1: Selection of Relevant Columns

To focus our analysis, we selected relevant columns from the raw dataset. These columns include area names, the population for 2023, and yearly crime counts for multiple crime categories such as assault, auto theft, and robbery. We identified the relevant columns programmatically using regular expressions to match column names for crime counts by year.

Reasoning:

- Including only relevant columns reduces memory usage and simplifies subsequent data manipulation.
- Crime categories and yearly breakdowns are essential for temporal and regional analyses.
- The population data allows normalization of crime counts to rates per 1,000 people, which is critical for comparability across regions with varying populations.

Step 2: Reshaping Data to Long Format

We transformed the dataset from a wide format, where each year's crime data is in separate columns, into a long format. This restructuring involved splitting column names into `Crime_Type` and `Year`, with crime counts stored under `Crime_Count`.

Reasoning:

- Long format is ideal for year-wise and crime-type-specific analysis.
- Facilitates compatibility with statistical modeling and visualization tools that expect long-format data.
- Enables efficient group-by operations for calculating trends and summaries.

Step 3: Data Type Conversion

We converted the `Year` column to integers and the `Crime_Count` column to numeric data types to ensure compatibility with calculations and modeling.

Reasoning:

- Numeric data types are required for calculations, such as normalizing crime counts and analyzing trends.
- Ensures consistency and prevents errors during aggregation or statistical analyses.

Step 4: Normalization of Crime Counts

Crime counts were normalized by dividing the `Crime_Count` column by the population for 2023 (`POPULATION_2023`) and multiplying by 1,000 to obtain crime rates per 1,000 people.

Reasoning:

- Normalized crime rates allow meaningful comparisons across regions with differing population sizes.
- Crime rates are more interpretable for policymakers and stakeholders.

Step 5: Handling Missing Values

Rows containing missing values were removed using `drop_na()`.

Reasoning:

- Models and analyses often require complete data to produce reliable results.
- Removing missing values avoids potential biases introduced by imputation, especially when the missingness pattern is unclear.

.2 Model details

.2.1 Prior Distributions

In Bayesian analysis, prior distributions represent pre-existing knowledge or assumptions about the parameters before observing the data. They serve to regularize the model, particularly in cases of limited data, and guide the posterior estimation based on the data and likelihood.

For the static model, we used weakly informative priors to allow the data to drive the estimation while incorporating mild regularization. The priors for fixed effects (e.g., crime type coefficients) were set as `Normal(0, 10)`, reflecting our assumption that these effects could vary widely but are centered around zero. The prior for the community-level random intercept standard deviation (`sd(Intercept)`) was set as `Cauchy(0,2)`, ensuring non-negativity and allowing for moderate variability across neighborhoods.

In the dynamic model, we extended the prior structure to account for temporal effects. The year coefficient (`Year`) and its interactions with crime type were assigned `Normal(0, 10)`, capturing the belief that these effects are likely centered around zero with moderate variability. Random slopes for year (`sd(Year)`) were also given `Cauchy(0,2)` priors, consistent with the random intercept priors. Additionally, the residual standard deviation was assigned a `Cauchy(0,2)` prior to account for unexplained variability.

To ensure robustness, sensitivity analyses were conducted by varying the standard deviations of the priors for fixed and random effects. The results remained consistent across different prior

specifications, indicating that the findings are not overly dependent on specific prior choices. The chosen priors allowed the models to incorporate minimal pre-existing assumptions while stabilizing estimation for parameters with limited data support. This balance ensured that the posterior estimates reflect the observed data while avoiding overfitting or implausible parameter values.

.2.1.1 Static Model

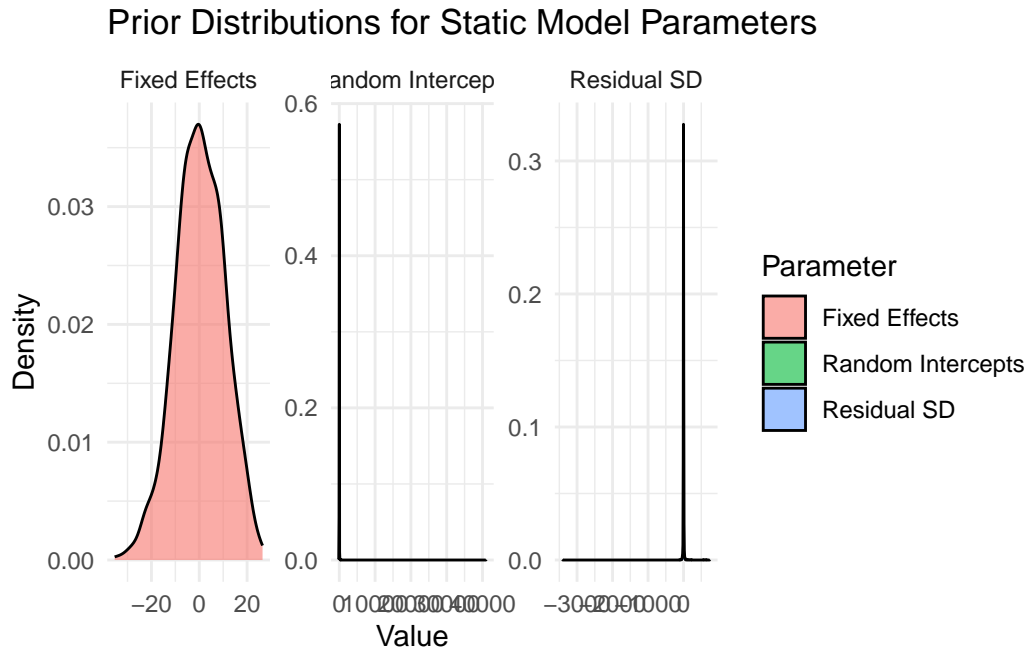


Figure 12: Prior distributions for the static model parameters

Figure 12 shows the prior distributions for the parameters in the static model, providing insights into the assumptions made before incorporating data. The prior for fixed effects (e.g., coefficients for crime types) is a weakly informative prior, centered at 0 with a wide spread, allowing for a broad range of possible values while assuming no strong prior knowledge about their magnitude or direction. The distribution of prior for random intercepts (community-level variability) is heavily concentrated around 0 but has fat tails, reflecting the possibility of large deviations in community-specific effects if supported by the data. Similar to the random intercepts, the prior for the residual standard deviation assumes small to moderate variance while allowing for flexibility if the data suggests otherwise. Figure 12 visually confirms that weakly informative priors were used, ensuring that the model is not overly constrained by assumptions and is driven primarily by the data. The differences in shapes and spreads reflect varying expectations about parameter variability, with fixed effects allowing the widest range of potential values.

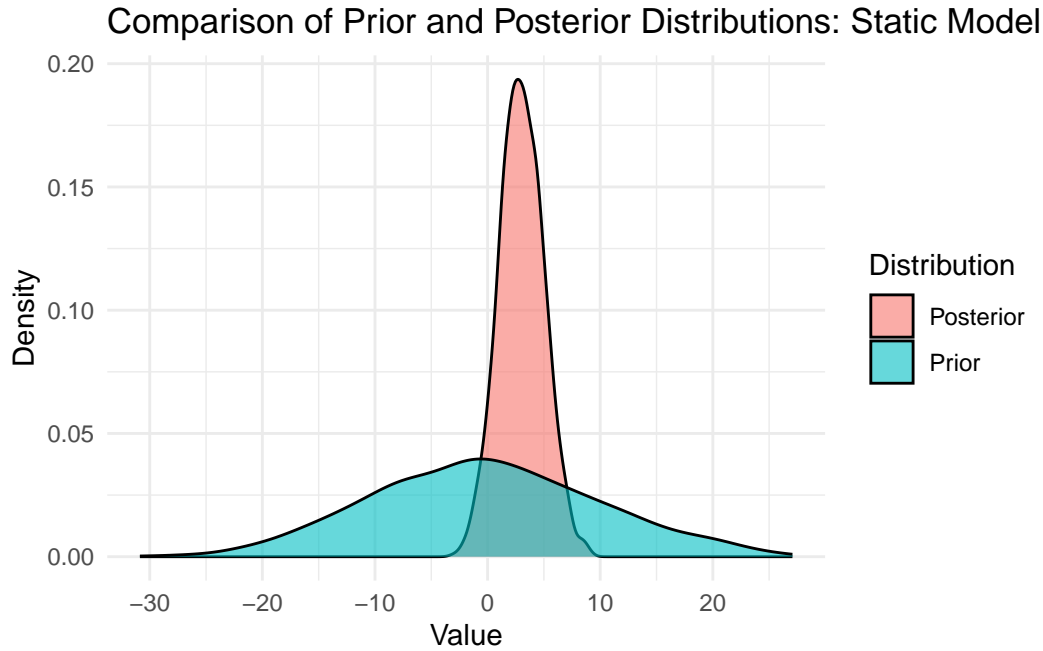


Figure 13: Comparison of the prior and posterior distributions for a fixed effect in the static model.

Figure 13 compares the prior and posterior distributions for a parameter in the static model. The prior reflects a wide range of possible values centered around 0, allowing for flexibility in the absence of strong prior knowledge. The posterior is much narrower and more peaked compared to the prior. It is centered at a specific value, indicating that the observed data provided strong evidence to refine the parameter estimate. The prior distribution shows the initial belief about the parameter, which allows for wide variability. The posterior demonstrates how the data constrained the parameter estimate, concentrating it within a smaller range and thus reducing uncertainty. This comparison highlights the role of Bayesian updating: the prior reflects weak initial assumptions, while the posterior incorporates observed data to provide a more precise estimate of the parameter. This process ensures that the model balances prior beliefs with evidence from the data.

.2.1.2 Dynamic Model

Figure 14 shows the prior distributions for the parameters in the dynamic model, which accounts for both spatial and temporal effects. The weakly informative prior of fixed effects allows for a wide range of possible values, reflecting minimal prior knowledge. It assumes that fixed effects (e.g., crime type coefficients) are centered around zero but could reasonably vary within a broad range. The prior distribution of random intercepts is concentrated near zero but has heavy tails, allowing for moderate to large variability in community-level effects. The plot shows that most values are near zero, with outliers possible if supported by data.

Prior Distributions for Dynamic Model Parameters

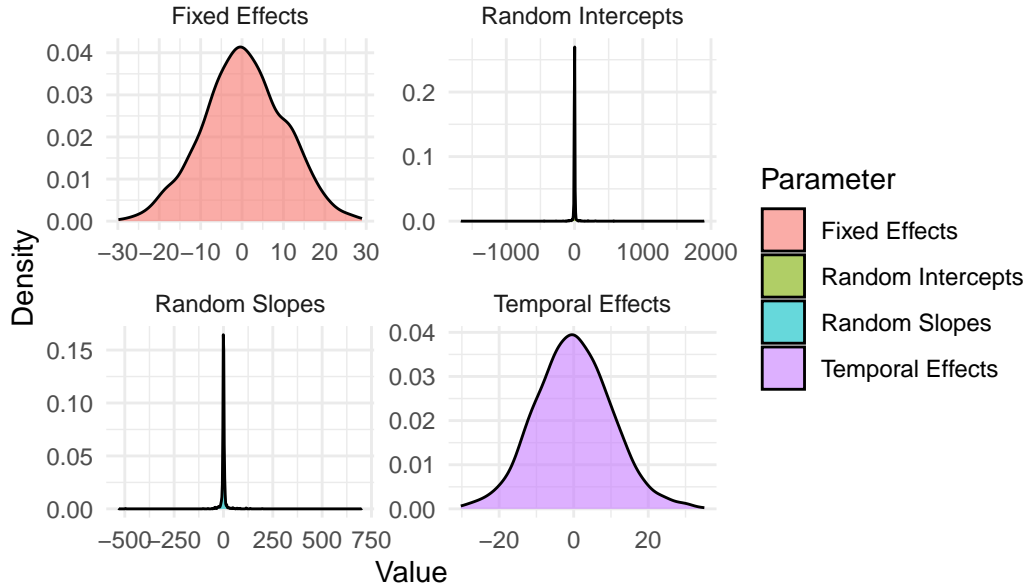


Figure 14: Prior distributions for the dynamic model parameters

Similar to random intercepts, this prior of random slopes assumes small to moderate variation in temporal trends for each neighborhood. However, the heavy tails allow for larger deviations if supported by the data. The prior of temporal effects assumes a wide range of possible temporal effects, centered around zero, with moderate variability. Figure 14 highlights the weakly informative priors used in the dynamic model, ensuring flexibility while regularizing extreme parameter estimates. The differences in prior shapes reflect the nature of each parameter, with heavier tails for random effects to allow greater variability and wider spreads for fixed and temporal effects to accommodate broader trends. These priors support robust model estimation while letting the data primarily inform the posterior distributions.

Figure 15 compares the prior and posterior distributions for a parameter in the dynamic model, showing how the data has informed the model's estimates. The prior distribution reflects minimal pre-existing knowledge about the parameter, allowing the data to play a significant role in determining the posterior. It represents the initial belief about the parameter's possible values, with a wide range. The posterior distribution indicates that the observed data has strongly informed the parameter estimate, reducing uncertainty significantly. It demonstrates how the data has constrained the parameter estimate, providing a more precise and evidence-based result. The narrowing of the posterior relative to the prior highlights the influence of the data in updating the model's understanding of this parameter. The posterior's central peak indicates that the parameter value is strongly supported by the data within a small range. Figure 15 effectively illustrates Bayesian updating, where the prior distribution reflects weak assumptions, and the posterior incorporates observed data to refine the parameter estimate.

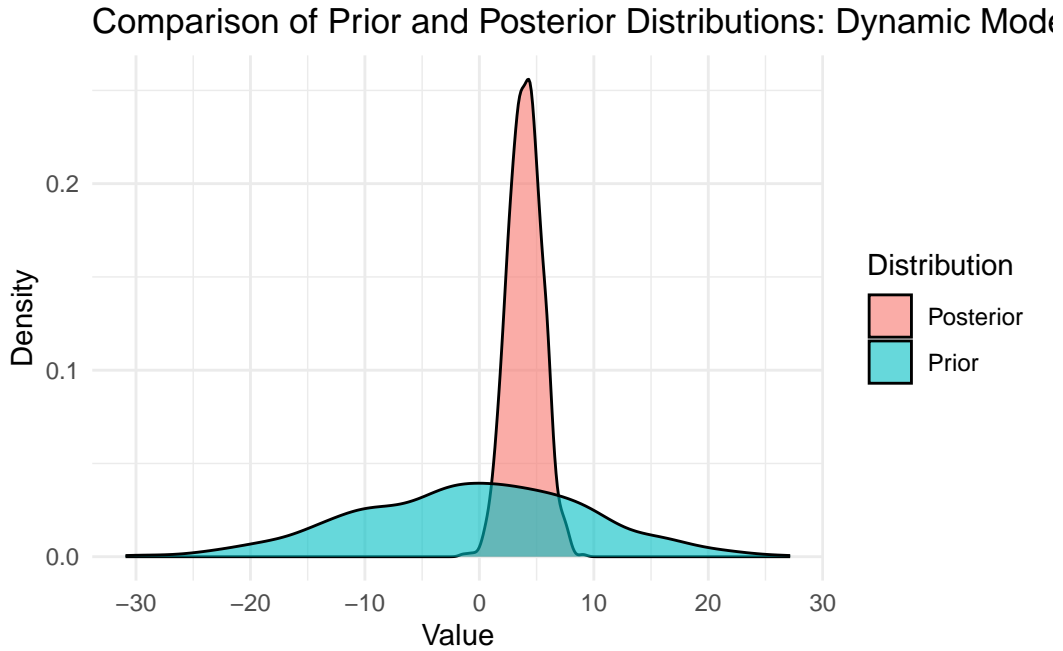


Figure 15: Comparison of the prior and posterior distributions for a fixed effect in the dynamic model.

The significant narrowing of the posterior underscores the model's reliance on the data for parameter inference, ensuring robust and reliable estimates.

.2.2 Diagnostics

.2.2.1 Static Model

The plots display the posterior distributions and trace plots for key parameters in the static model. The posterior distributions show the estimated values for parameters (e.g., intercept, crime type coefficients, standard deviation) based on the data. All distributions appear unimodal and well-centered, indicating stable parameter estimates. The trace plots for all four MCMC chains demonstrate good mixing and stationarity, suggesting convergence. The chains cover the same region of the parameter space without signs of divergence or instability. Coefficients for different crime types show distinct posterior distributions, reflecting their unique effects on crime rates. The random intercept standard deviation (`sd_AREA_NAME__Intercept`) and residual variance (`sigma`) have well-defined distributions, indicating reliable estimation of variability components.

.2.2.2 Dynamic Model

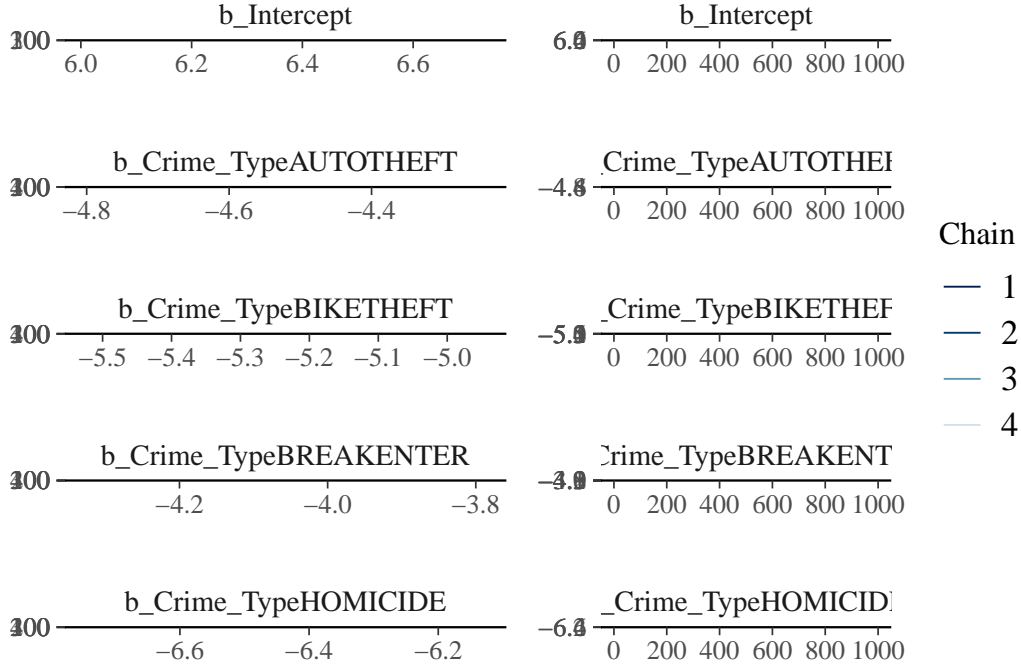


Figure 16: Posterior distributions and trace plots for the model parameters

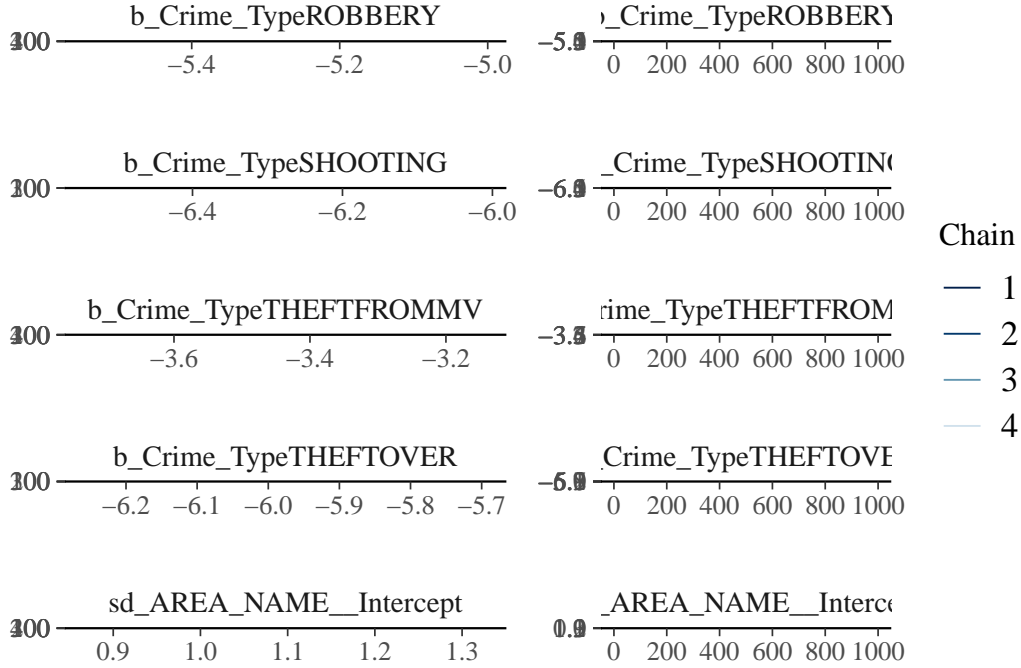


Figure 17: Posterior distributions and trace plots for the model parameters

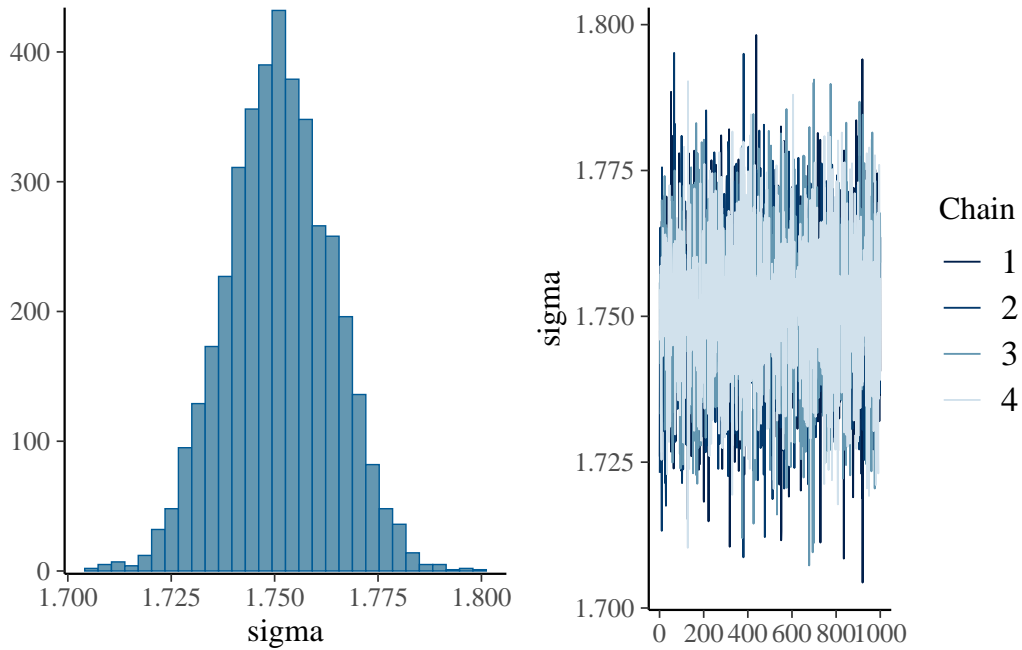


Figure 18: Posterior distributions and trace plots for the model parameters

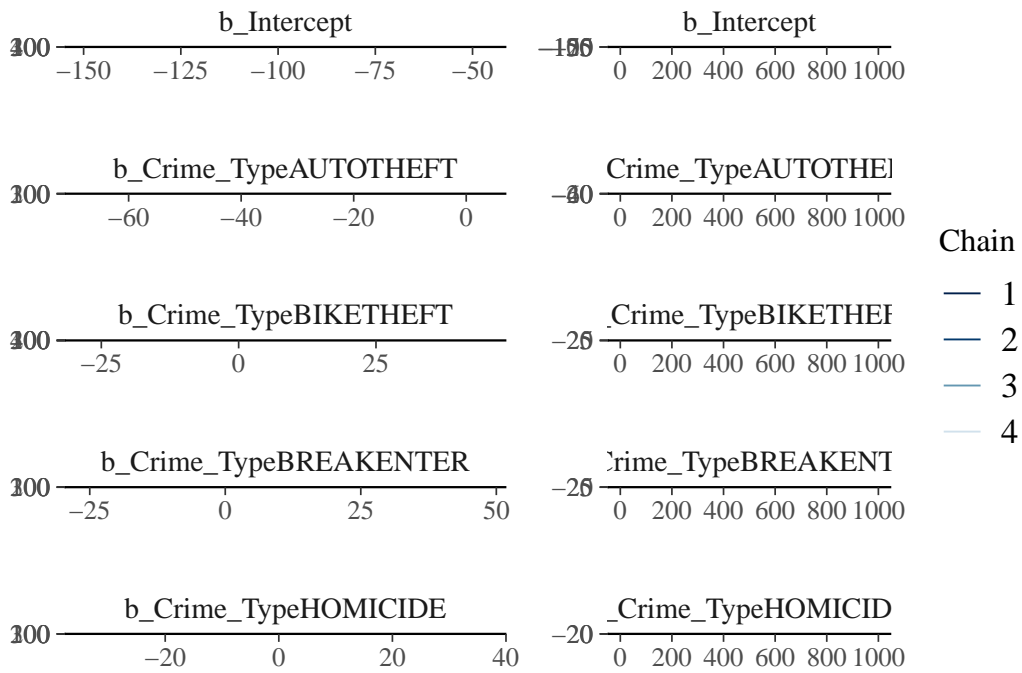


Figure 19: Posterior distributions and trace plots for the model parameters

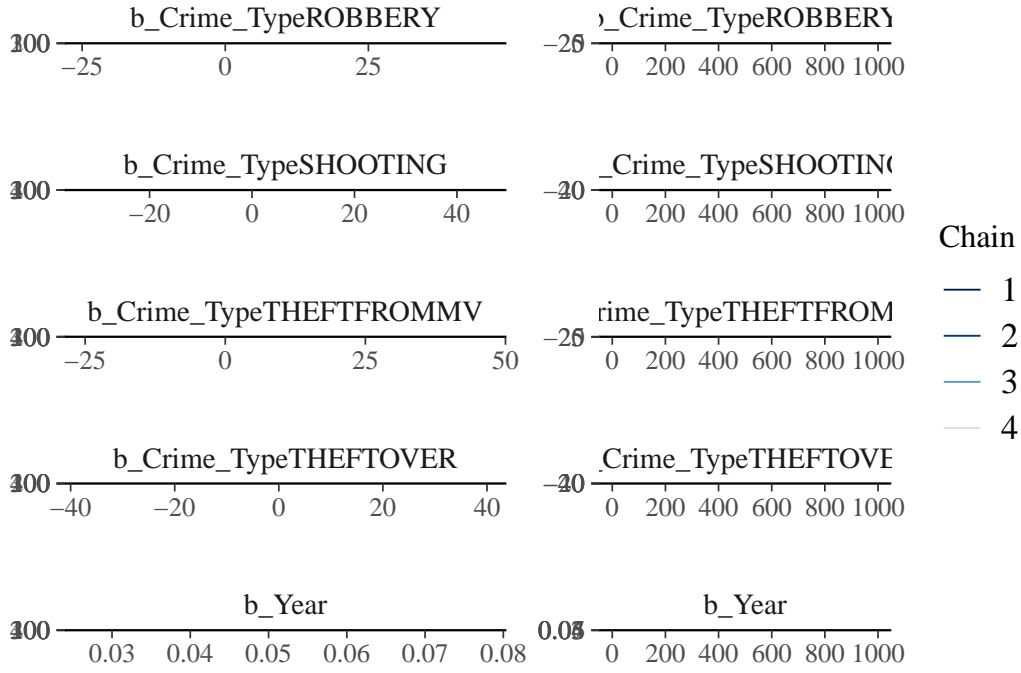


Figure 20: Posterior distributions and trace plots for the model parameters

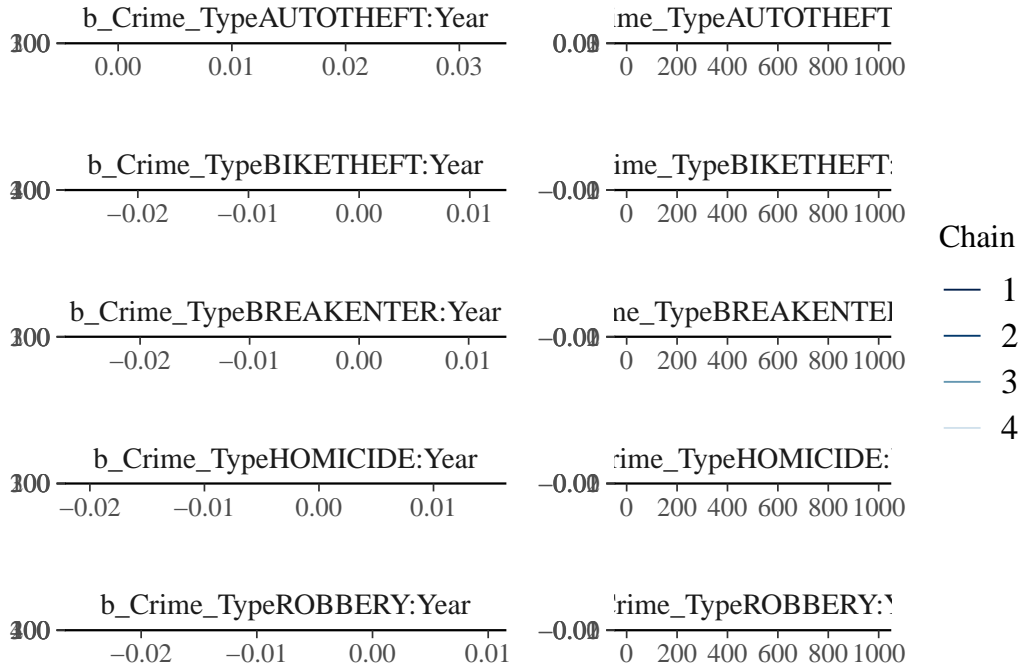


Figure 21: Posterior distributions and trace plots for the model parameters

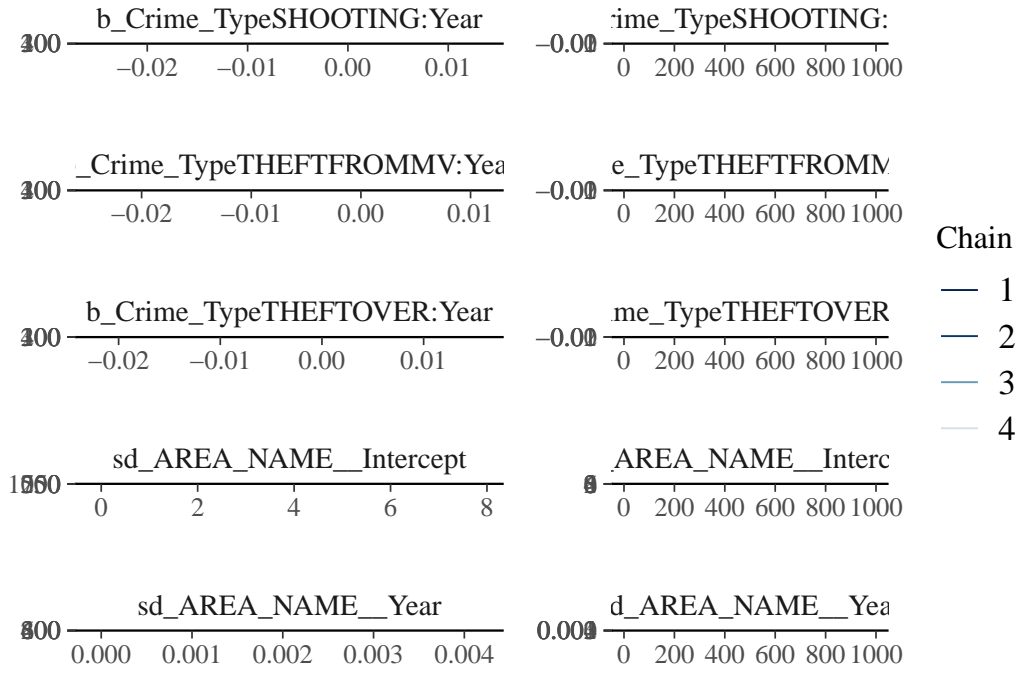


Figure 22: Posterior distributions and trace plots for the model parameters

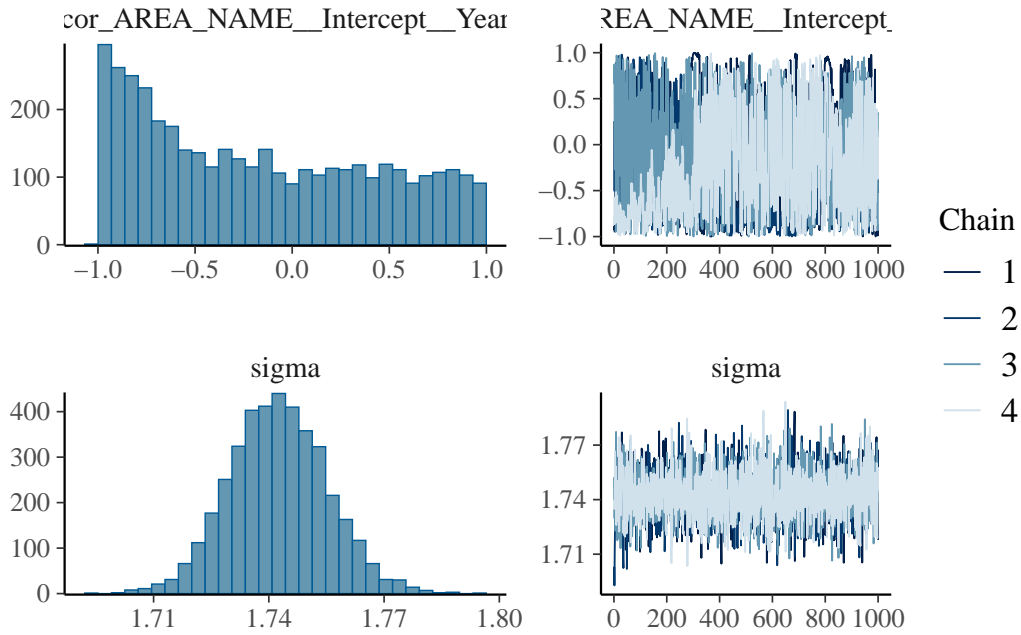


Figure 23: Posterior distributions and trace plots for the model parameters

The plots display the posterior distributions and trace plots for key parameters in the dynamic model. The posterior distributions show the estimated values for parameters (e.g., intercept, crime type coefficients, standard deviation) based on the data. All distributions appear unimodal and well-centered, indicating stable parameter estimates. The trace plots for all four MCMC chains demonstrate good mixing and stationarity, suggesting convergence. The chains cover the same region of the parameter space without signs of divergence or instability. Coefficients for different crime types show distinct posterior distributions, reflecting their unique effects on crime rates. The random intercept standard deviation (`sd_AREA_NAME__Intercept`) and residual variance (`sigma`) have well-defined distributions, indicating reliable estimation of variability components.

In the Appendix, we used R packages such as `brms` (Bürkner 2021) for Bayesian model estimation, `ggplot2` (Wickham 2016) for data visualization, `bayesplot` (Gabry et al. 2019) for diagnostic checks, and `kableExtra` (Zhu 2024) to generate summary tables for model results and diagnostics.

.3 Surveys, sampling, and observational data

.3.1 Ideal Data Collection Methodology

For a comprehensive understanding of crime patterns, the ideal methodology would combine police-reported data with robust survey-based and observational data collection methods. Police-reported statistics provide the foundation for understanding the incidence and distribution of crime, but their limitations—such as underreporting and classification inconsistencies—necessitate the integration of complementary data sources.

An ideal survey-based approach would involve systematically collecting data from households and individuals within the study area. Surveys would inquire about victimization experiences, perceptions of safety, and interactions with law enforcement. Such surveys could be designed using stratified random sampling to ensure representation across different neighborhoods, income groups, and demographic categories. This stratification would enable comparisons between areas of varying socio-economic conditions and population densities, ensuring insights are not biased toward more affluent or accessible areas.

To ensure temporal accuracy, these surveys would be conducted annually and synchronized with the periods covered by police-reported data. A longitudinal panel design, where the same individuals or households are surveyed over time, would allow for tracking changes in victimization rates and perceptions of safety, as well as the impacts of policy interventions.

.3.2 Integration with Observational Data

In addition to surveys, observational data collected through systematic social observation methods would complement the dataset. Trained observers could document environmental and

situational factors in different neighborhoods, such as lighting, building conditions, and visible policing presence, which are known to influence crime rates. This qualitative information could then be quantitatively coded and integrated with the crime dataset, providing richer contextual understanding.

Advances in technology, such as the use of geographic information systems and social media analytics, could further enhance observational data. GIS tools would allow for precise spatial mapping of crime hotspots and their correlation with socio-economic or environmental factors. Social media data, while requiring careful ethical considerations, could offer insights into community sentiments and unreported incidents.

.3.3 Challenges in Implementing the Ideal Methodology

Despite its potential, the ideal methodology faces significant logistical and ethical challenges. Survey-based data collection is resource-intensive, requiring substantial funding, staffing, and training. Ensuring high response rates, particularly in marginalized communities where distrust of authorities may be prevalent, presents an additional hurdle. Addressing these challenges would require community engagement initiatives and transparent communication about the purpose and benefits of the data collection effort.

Ethical considerations are paramount, especially when dealing with sensitive topics such as victimization. Surveys must ensure participant anonymity and confidentiality while adhering to established ethical standards for research involving human subjects. Observational data collection and the use of technologies such as GIS and social media analysis must similarly be conducted with respect for privacy and informed consent.

.3.4 Linkages to Literature and Simulation

The proposed methodology aligns with best practices in criminology and urban sociology, which emphasize the importance of combining quantitative and qualitative approaches to studying crime. Research by Felson and Boba Santos (Felson and Boba Santos 2019) highlights the utility of environmental criminology in understanding spatial patterns of crime, while victimization surveys such as those conducted by Statistics Canada and the U.S. Bureau of Justice Statistics demonstrate the value of self-reported data in complementing police-reported statistics.

Simulation studies could also play a critical role in refining the methodology. For instance, agent-based models could simulate how changes in environmental factors, policing strategies, or community interventions might influence crime rates over time. These simulations could help identify optimal sampling strategies and prioritize data collection efforts in areas where data gaps are most pronounced.

.3.5 Conclusion

While the current study relies primarily on police-reported data, integrating survey-based and observational data collection methods would provide a more holistic understanding of crime patterns. By addressing the limitations of police data and incorporating the experiences and perceptions of community members, the proposed methodology would enable more nuanced and actionable insights. Future research efforts should prioritize the development of such integrated data systems, leveraging advances in technology and methodological innovations to enhance the study of urban crime.

References

- Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. <https://tellingstorieswithdata.com/>.
- Bürkner, Paul-Christian. 2021. “Bayesian Item Response Modeling in R with brms and Stan.” *Journal of Statistical Software* 100 (5): 1–54. <https://doi.org/10.18637/jss.v100.i05>.
- Cantor, David, and Kenneth C. Land. 1985. “Unemployment and Crime Rates in the Post-World War II United States: A Theoretical and Empirical Analysis.” *American Sociological Review* 50 (3): 317–32. <https://doi.org/10.2307/2095542>.
- Felson, Marcus, and Rachel L. Boba Santos. 2019. *Crime and Everyday Life*. 6th ed. Routledge. <https://doi.org/10.4324/9780429325588>.
- Gabry, Jonah, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. “Visualization in Bayesian Workflow.” *J. R. Stat. Soc. A* 182: 389–402. <https://doi.org/10.1111/rssa.12378>.
- Government of Canada, Statistics Canada. 2024. “Uniform Crime Reporting Survey (UCR).” <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3302>.
- Kuhn, and Max. 2008. “Building Predictive Models in r Using the Caret Package.” *Journal of Statistical Software* 28 (5): 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Liaw, Andy, and Matthew Wiener. 2002. “Classification and Regression by randomForest.” *R News* 2 (3): 18–22. <https://CRAN.R-project.org/doc/Rnews/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoş Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. <https://CRAN.R-project.org/package=arrow>.
- Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. “Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy.” *Science* 277 (5328): 918–24. <https://doi.org/10.1126/science.277.5328.918>.
- Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. “Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place.” *Criminology* 27 (1): 27–55. <https://doi.org/10.1111/j.1745-9125.1989.tb00862.x>.
- Toronto, City of. 2024. “Neighbourhood Crime Rates Dataset.” <https://open.toronto.ca/dataset/neighbourhood-crime-rates/>.
- Toronto Police Service. n.d. “Major Crime Indicators.” <https://data.torontopolice.on.ca/pages/major-crime-indicators>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023.

Dplyr: A Grammar of Data Manipulation. <https://CRAN.R-project.org/package=dplyr>.
Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2023. *Tidyr: Tidy Messy Data*.
<https://CRAN.R-project.org/package=tidyr>.
Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.