

# The Association of Crime Occurrence on Rates of Suspensions and Expulsions in Chicago Public Schools Using Double Hurdle Models

Nicholas Fox, John Cote, Dr. Earvin Balderama  
Department of Mathematics & Statistics, Loyola University Chicago, Chicago, IL, USA



## Motivation

- **Goal:**
  - ▷ To gain a better understanding of factors that may influence student expulsion
- **Steps:**
  - ▷ Investigate potential factors leading to student expulsion
  - ▷ Analyze the effects of the amount of violent crime in a school attendance boundary
  - ▷ Examine spatial effects of school attendance boundaries

## Model

### Priors for Beta and Alpha Parameters

$$\beta_j \sim \text{Normal}(0, 100) \quad (1)$$
$$\alpha_i \sim \text{Normal}(0, 1000)$$

### Negative Binomial and Generalized Pareto Distribution Double Hurdles Models

**p** = P(zero-count)

$$\text{logit}(\mathbf{p}) = \mathbf{X}\gamma + \mathbf{V}\alpha$$

**μ** = mean of typical-count distribution.

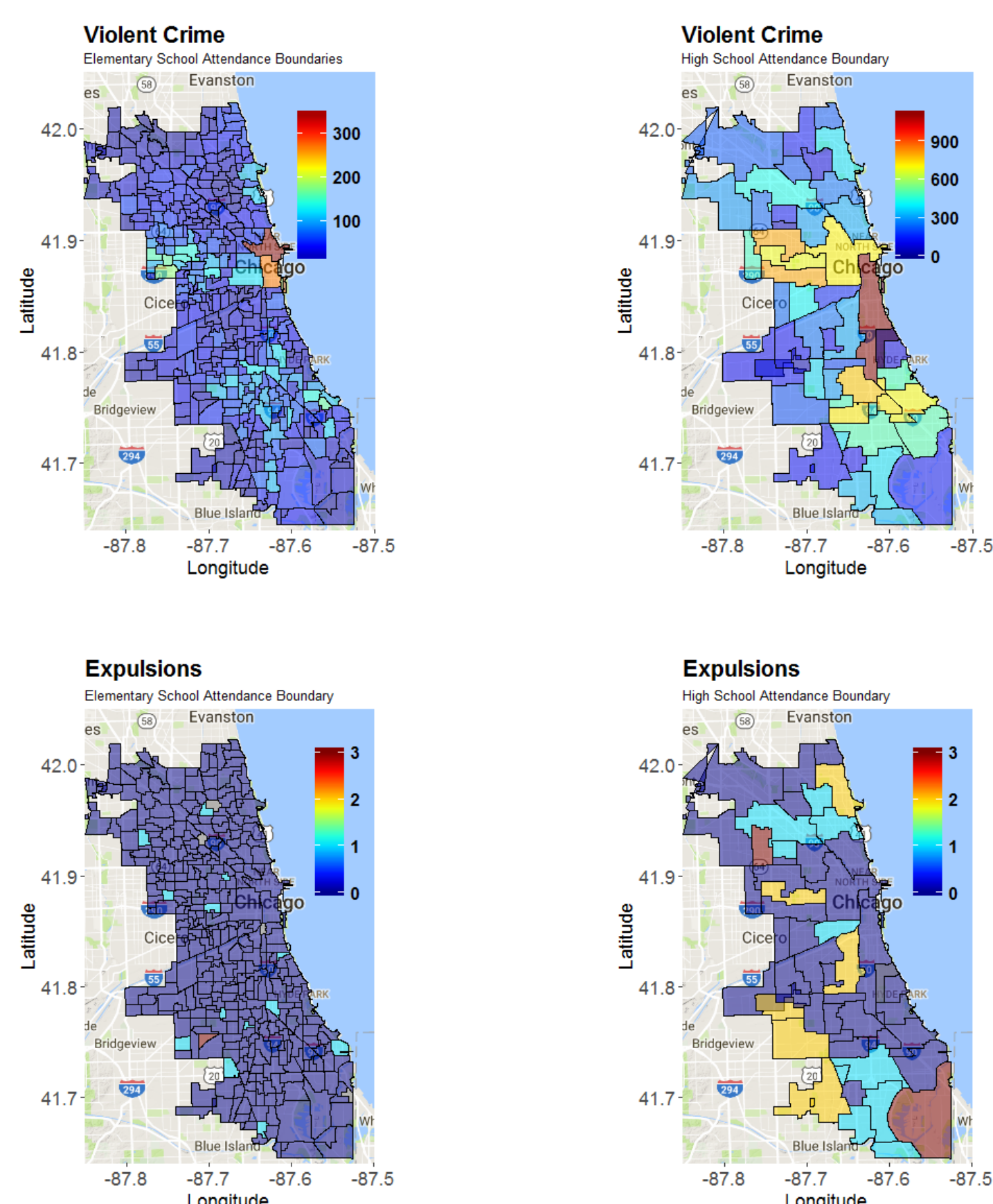
$$\text{log}(\boldsymbol{\mu}) = \text{log}(\text{School.Pop}) + \mathbf{X}\beta + \mathbf{V}\alpha$$

**q** = P(large-count | nonzero-count)

$$\text{logit}(\mathbf{q}) = \mathbf{X}\delta$$

## Data

- Chicago Data Portal
  - ▷ Chicago Crime Records for the year 2015
  - ▷ Shapefiles for CPS attendance boundaries
- Chicago Public School Data from the CPS Website
  - ▷ Attendance records
  - ▷ Education quality scores
  - ▷ Standardized test scores
  - ▷ Suspension, expulsion, and misconduct information



## Model Variables and Results

Response data:

Expulsions → Expulsions counts for each Chicago Public School  
Suspensions → Suspension counts for each Chicago Public School

Model Covariates covariates:

- Common Covariates
  - $\mathbf{x}_1 - \mathbf{x}_3$  = Ethnicity counts for schools; White, African American, Hispanic respectively
  - $\mathbf{x}_4$  = Number of misconducts per school
- Unique Elementary School Covariates
  - $\mathbf{x}_5$  = School quality score
  - $\mathbf{x}_6$  = PARCC average math score
  - $\mathbf{x}_7$  = PARCC English Language Arts average score
  - $\mathbf{x}_8$  = Crime counts based on school attendance boundary
- Unique High School Covariates
  - $\mathbf{x}_5$  = Average ACT composite score per school
  - $\mathbf{x}_6$  = Crime counts based on school attendance boundary
  - $\mathbf{x}_7$  = School quality score
- Spatial Random Effects are included for each model

### Crime Regression Coefficients for Elementary School Models

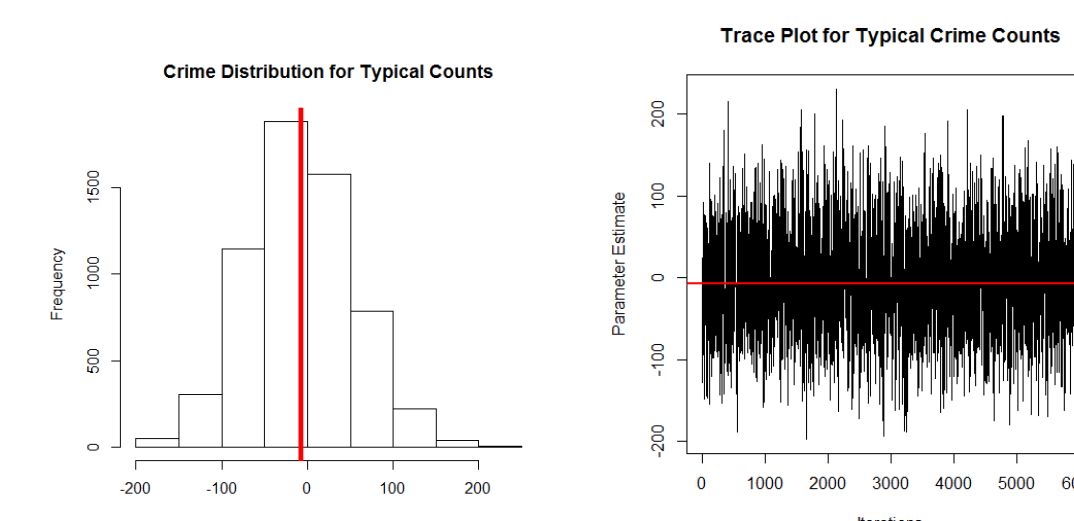


Figure: Expulsion Models for Elementary Schools

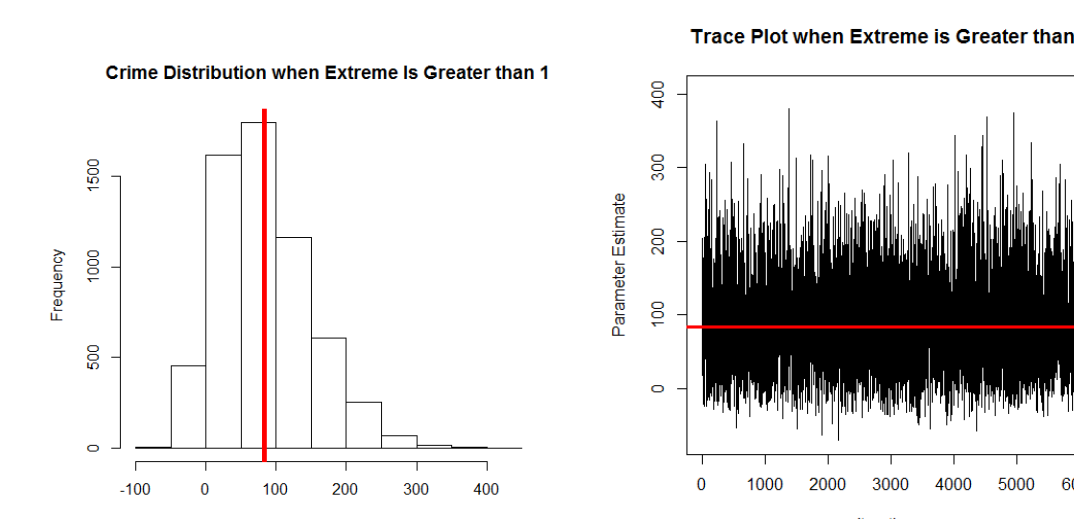


Figure: Suspension Models for Elementary Schools

### Crime Regression Coefficients for High School Models

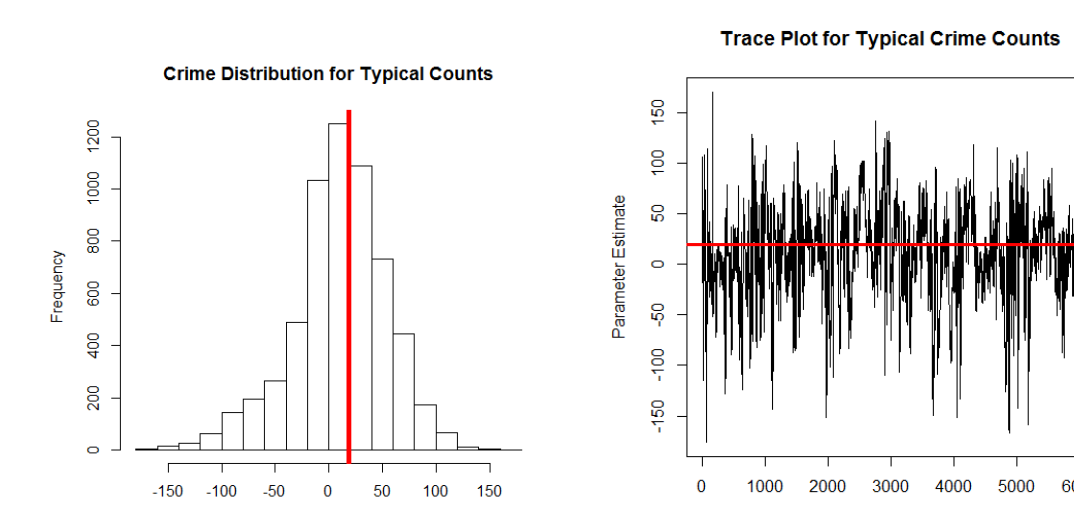


Figure: Expulsion Models for High Schools

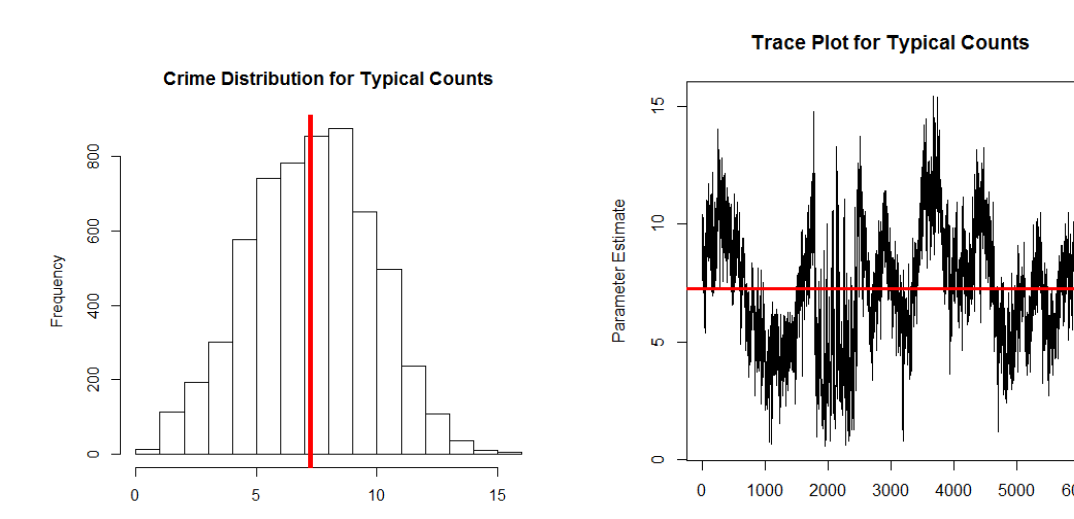


Figure: Suspension Models for High Schools

- **Interpretation Example:** A parameter estimate of 7.25 has a 1408 multiplicative effect on the mean number of suspensions

## Model Diagnostics

| Model   | DIC      | CPO      | PPO      |
|---------|----------|----------|----------|
| 1       | 189.0624 | 0.906428 | 0.910272 |
| 95      | 186.6292 | 0.906297 | 0.909889 |
| Extreme | 187.1526 | 0.90624  | 0.910572 |

| Model   | DIC      | CPO      | PPO      |
|---------|----------|----------|----------|
| 1       | 3891.406 | 0.009427 | 0.009448 |
| 95      | 3901.314 | 0.008998 | 0.009682 |
| Extreme | 3891.117 | 0.009428 | 0.009446 |

Figure: Elementary School Expulsion Diagnostic (left) and Suspension Diagnostics (right)

| Model   | DIC      | CPO      | PPO      |
|---------|----------|----------|----------|
| 1       | 399.6648 | 0.425075 | 0.445602 |
| 95      | 370.6124 | 0.422647 | 0.449454 |
| Extreme | 371.3801 | 0.430773 | 0.455491 |

| Model   | DIC      | CPO      | PPO      |
|---------|----------|----------|----------|
| 1       | 1140.971 | 0.010367 | 0.010546 |
| 95      | 1045.549 | 0.045831 | 0.060184 |
| Extreme | 1135.263 | 0.015739 | 0.01652  |

Figure: High School Expulsion Diagnostic (left) and Suspension Diagnostics (right)

## Discussion of Results

- Demographics
  - ▷ These factors were mostly insignificant
  - ▷ A couple were barely significant in a couple models, however overall not very important
- Test Scores
  - ▷ Test scores were used to look at student performance and were used as a proxy for student motivation
  - ▷ Most significant when modelling the mean number of non-zero expulsion counts
- Misconducts
  - ▷ The number of misconducts was significant in the majority of the models run. This is expected, as with the more misconducts we expect more suspensions and expulsions
- Quality Scores
  - ▷ A quality score provided by CPS for each school
  - ▷ Found to be insignificant in the majority of the models
- Violent Crime Frequency
  - ▷ Crime frequency was found to be significant for both suspensions and expulsions for most of the models
  - ▷ For most cases, it was found to be a positive relationship between the amount violent crime and the number of suspension and expulsions
- Spatial Random Effects
  - ▷ This variable examined whether the occurrence of a suspension or expulsion at a school affects a neighboring school
  - ▷ Found to be almost completely insignificant, implying there are not spatial influences between schools

## Future Considerations

- Consider temporal factors as potentially affecting suspensions and expulsions
- Consider more types of crime than just violent crime
- Consider physical distances of crimes to a school

## Acknowledgements

- Software Used: R (www.r-project.org)
- Data obtained from the Chicago Data Portal and Chicago Public Schools Data Page
- Special thanks to the Loyola University Chicago Department of Mathematics and Statistics