Multi-Moment Spatial Analysis of Violence

Avinash S. Bhati, Ph.D.
Senior Research Associate

March 28—31, 2007

Presentation at the 9th Crime Mapping Research Conference, Omni William Penn Hotel, Pittsburgh, PA.

URBAN INSTITUTE
Justice Policy Center

Introduction and background

- Count Outcome Examples:
  - Number of Job Changes, Ship Accidents, Strikes, Patents Filed, Crimes Recorded, etc.
- Count Outcome/Model Peculiarities:
  - Low non-negative integers (0,1,2,3,…)
  - Discrete/Skewed
  - May exhibit extra-Poisson variation
  - Preponderance of one or more counts (e.g., 0)
- Correlation Across Space
  - Can be substantive or a nuisance
  - Usually contaminates inferences
  - Particularly difficult to deal with in non-linear models

Overview of this presentation

- What are moments of a distribution?
- Derive a Multi-Moment Generalized Poisson Model:
  - Amount of violence
  - Predictability of violence
  - Mere occurrence of a violent event
  - Substantive spatial structure
- Simulated evidence
- Homicides in Chicago neighborhoods
- Model implications
- Discussion

Distributional Moments

- If x_n are a sample of some random variable, then its
  - r[th] *raw* moments is defined as: $\kappa_r = \frac{1}{N}\sum_n x_n^r$
  - r[th] *central* moments is defined as: $\mu_r = \frac{1}{N}\sum_n (x_n - \mu_1)^r$
  - r[th] *geometric* moments is defined as: $\gamma_r = \frac{1}{N}\sum_n (\log x_n)^r$
- Generically, we can define moments as the mean of various transformations of the original variable x_n

$$\eta_r = \frac{1}{N}\sum_n \psi_r(x_n)$$

Why Study Multiple Moments?

- Typically researchers study only the 1^{st} moment: the expected value. But what about …
  - the expected variance, skewness, etc.
- Example from Financial Econometrics (returns on investment). Investors prefer …
  - Higher returns (i.e., higher mean return)
  - More certain return (i.e., more predictable returns)
  - Left Skewed returns (i.e., surprise gain better than surprise loss)
- Different moments capture different aspects of the phenomenon under study

Setting up the generic problem - I

- *N* Observed Outcomes: $\mathbf{y} = (y_1, \ldots, y_N)$
- *N* Emitted Signals: $\mathbf{s} = (s_1, \ldots, s_N)$
- Approximate relationship at the individual level

$$y_n \approx s_n \qquad \forall n$$

- Partial guidance from theory: *the presence or magnitude of some attributes induce variation in the signals.* I.e., suggests relevance of attributes.

Qualitative Relevance

$$0 \neq \sum_n x_{kn} s_n$$

Quantitative Relevance

$$\sum_n x_{kn} y_n = \sum_n x_{kn} s_n$$

Setting up the generic problem - II

- *Step 1*: Re-parameterize unknown signals into proper probabilities (as weighted sum of propositions)

$$s_n = \mathbf{z}'\mathbf{p}_n = \sum_m z_m p_{mn} \qquad \sum_m p_{mn} = 1, \forall n \qquad (1)$$

- *Step 2*: Quantitative relevance of attributes yield constraints on the probabilities.

$$\sum_n x_{kn} y_n = \sum_n x_{kn} \mathbf{z}'\mathbf{p}_n \qquad \forall k \qquad (2)$$

- *Result*: Ill-posed inversion problem (infinite solutions)

Setting up the generic problem - III

- *IT Solution*: Choose most uncertain / least informative solution using Maximum Entropy (Ed Jaynes)

$$\max_{\mathbf{p}} H = -\sum_n \mathbf{p}_n' \log \mathbf{p}_n \quad \text{s.t. (1) \& (2)}$$

- If corresponding priors ($\mathbf{p}^0$) exist, choose solution that minimizes Cross Entropy

$$\min_{\mathbf{p}} CE = \sum_n \mathbf{p}_n' \log(\mathbf{p}_n / \mathbf{p}_n^0) \quad \text{s.t. (1) \& (2)}$$

- If priors are uniform, both problems are identical

Setting up the generic problem - IV

- Primal *Constrained* Optimization Problem

$$L = \sum_n \mathbf{p}_n' \log(\mathbf{p}_n / \mathbf{p}_n^0) + \sum_n \eta_n (1 - \mathbf{1}'\mathbf{p}_n) + \sum_k \lambda_k \left(\sum_n x_{kn} y_n - \sum_n x_{kn} \mathbf{z}'\mathbf{p}_n \right)$$

- Optimal Solution

$$p_{mn} = \frac{p_m^0 \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})}{\sum_m p_m^0 \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})} = \frac{p_m^0 \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})}{\Omega_n}$$

- Resulting Dual *Unconstrained* Optimization Problem

$$L_* = \sum_{kn} x_{kn} y_n \lambda_k - \sum_n \log \boldsymbol{\Omega}_n$$

The Poisson model derived

- Define Support Space as: $\mathbf{z} = (0,1,2,\ldots,z_M)$
- Assume Priors are: $p_m^0 = {}^{1}\!/_{z_m!}$
- The solution is …

$$p_{mn} = \frac{(1/z_m!)\exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})}{\sum_m (1/z_m!)\exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})} = \frac{(1/z_m!)\exp(z_m \mathbf{x}_n' \boldsymbol{\lambda})}{\exp(\exp(\mathbf{x}_n' \boldsymbol{\lambda}))}$$

$$= \frac{\exp(-\exp(\mathbf{x}_n' \boldsymbol{\lambda}))\exp(\mathbf{x}_n' \boldsymbol{\lambda})^{z_m}}{z_m!} = \frac{\exp(-\alpha_n)\alpha_n^{\ z_m}}{z_m!}$$

- … the Poisson model with a log-link function
- But, the Poisson model could be mis-specified …

Generalized Poisson Model - I

- … potentially mis-specified Poisson model means potentially incorrect priors $1/z_m!$.
- Solution: Parameterize dependence of p_{mn} on priors.
- Re-write solution as:

$$\begin{aligned} p_{mn} &= \boldsymbol{\Omega}^{-1} z_m!^{-1} \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda}) \\ &= \Omega^{-1} z_m!^{-1+\delta} \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda}) \\ &= \Omega^{-1} z_m!^{-1} \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda} + \log z_m! \delta) \end{aligned}$$

- Generalized Poisson: replace δ by $\mathbf{x}_n' \boldsymbol{\beta}$.

$$p_{mn} = \Omega^{-1} z_m!^{-1} \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda} + \log z_m! \mathbf{x}_n' \boldsymbol{\beta})$$

Generalized Poisson Model - II

- How to obtain this generalized solution?
- Simultaneously impose constraints

$$\sum_n x_{kn} y_n = \sum_{nm} x_{kn} z_m p_{mn} \qquad \forall k$$

$$\sum_n x_{kn} \log(y_n !) = \sum_{nm} x_{kn} \log(z_m !) p_{mn} \qquad \forall k$$

- Which yields the desired solution

$$p_{mn} = \Omega^{-1} z_m !^{-1} \exp(z_m \mathbf{x}_n' \boldsymbol{\lambda} + \log z_m ! \mathbf{x}_n' \boldsymbol{\beta})$$

- More generally, impose multi-moment constraints

$$\sum_n x_{kn} \psi_j (y_n) = \sum_{nm} x_{kn} \psi_j (z_m) p_{mn} \qquad \forall k, j$$

Spatial Structure in the Outcomes

- Spatial Autocorrelation: Standard Poisson process

$$\left.\begin{aligned} \log \boldsymbol{\alpha} &= \mathbf{X}\boldsymbol{\lambda} + \rho \mathbf{W} \log \boldsymbol{\alpha} \\ \log \boldsymbol{\alpha} &= (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}\boldsymbol{\lambda} \\ \log \boldsymbol{\alpha} &= \tilde{\mathbf{X}}\boldsymbol{\lambda} \end{aligned}\right\} \quad \mathbf{y} \sim \text{Poisson}(\boldsymbol{\alpha})$$

- Modified constraints needed

$$\sum_{n} \tilde{x}_{kn} y_n = \sum_{nm} \tilde{x}_{kn} z_m p_{nm}$$

- Generalized Poisson process with spatial autocorrelation among multiple moments

$$\sum_{n} \tilde{x}_{jkn} \psi_j(y_n) = \sum_{nm} \tilde{x}_{jkn} \psi_j(z_m) p_{nm}$$

Testing Hypothesis and Specifications

- Asymptotic Covariance of parameters: approximated by the negative inverted Hessian of the dual objective function:

$$\mathrm{Cov}(\boldsymbol{\theta}) = \left\{ -\frac{\partial^2 L_*}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right\}^{-1} \quad \text{where} \quad \boldsymbol{\theta} = (\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\rho})'$$

- Nested Specifications: Entropy Ratio Statistic

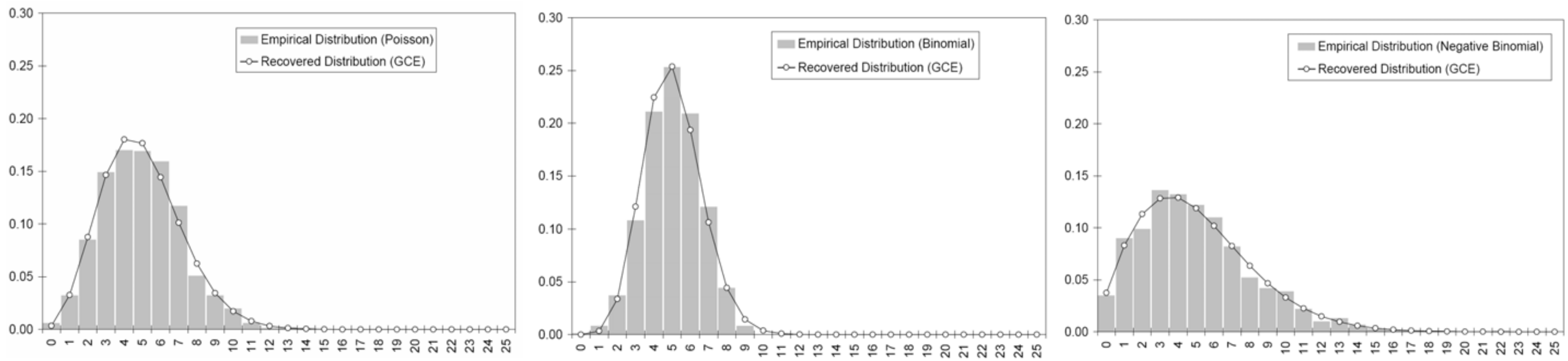
$$ER = 2 \times \left(L_{*(U)} - L_{*(R)} \right) \rightarrow \chi^2_{(R)}$$

- Can always use Huber-White Robust Standard Errors (is we suspect remaining structure in errors)

Yeah! But does it really work …

A Single Randomly Generated Sample

- Does the Generalized Poisson Model recover over-dispersion or under-dispersion in the data?



Some Simulation Results

1. How does the *ER* statistic used for testing unobserved heterogeneity perform?
2. Consequences of ignoring unobserved heterogeneity (with and without nuisance spatial autocorrelation in this heterogeneity)
   - Over rejection under the null, i.e., $\lambda_k = 0$ or $\rho = 0$
   - Does the Generalized Poisson model do better?

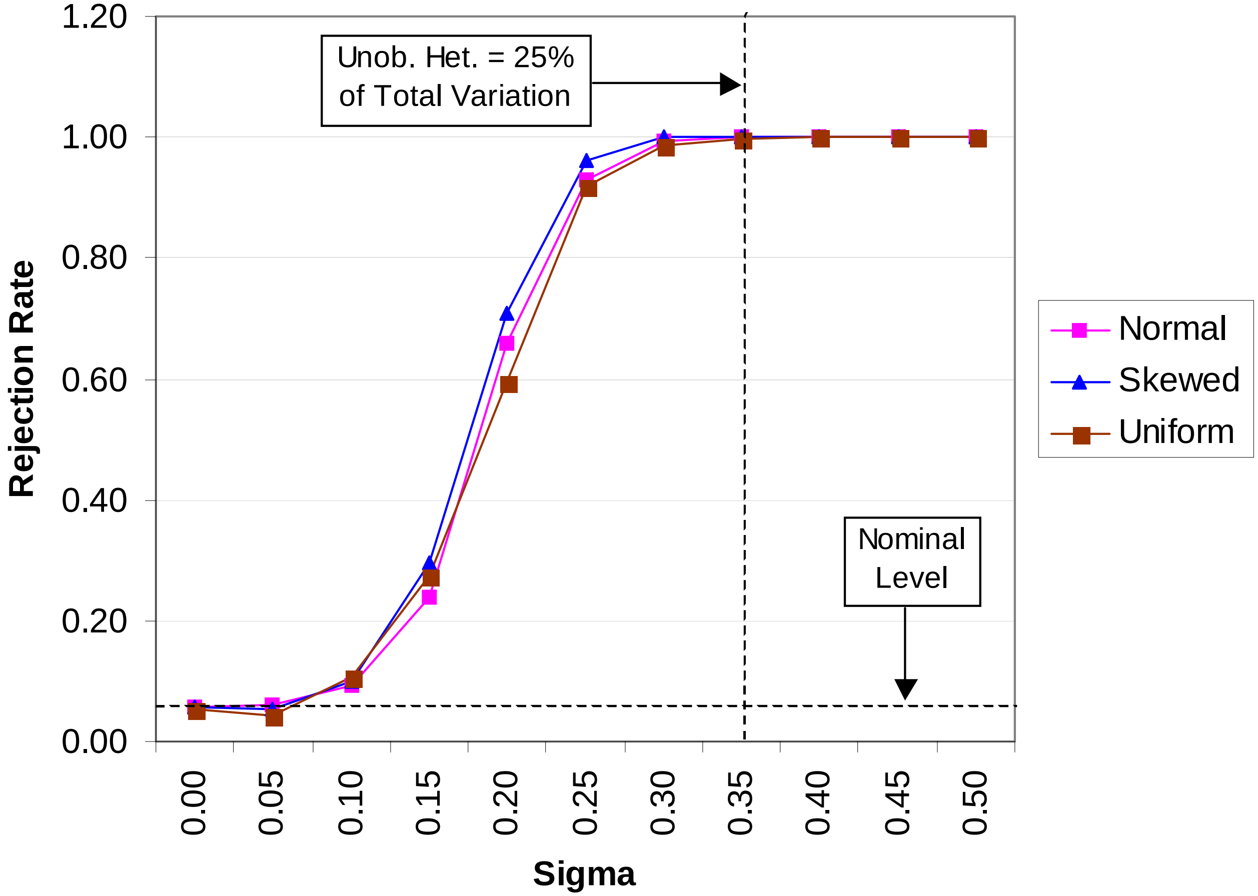
- Simulation design: $y_n \sim \text{Po}(\alpha_n)$ *N*=343

where $$\log \alpha_n = \underbrace{\beta_0 + \beta_1 \times N(0,1)}_{\text{fixed across repeated samples}} + e_n \qquad \beta_0 = \beta_1 = 1$$

with varying form and structure in e_n across 500 reps.

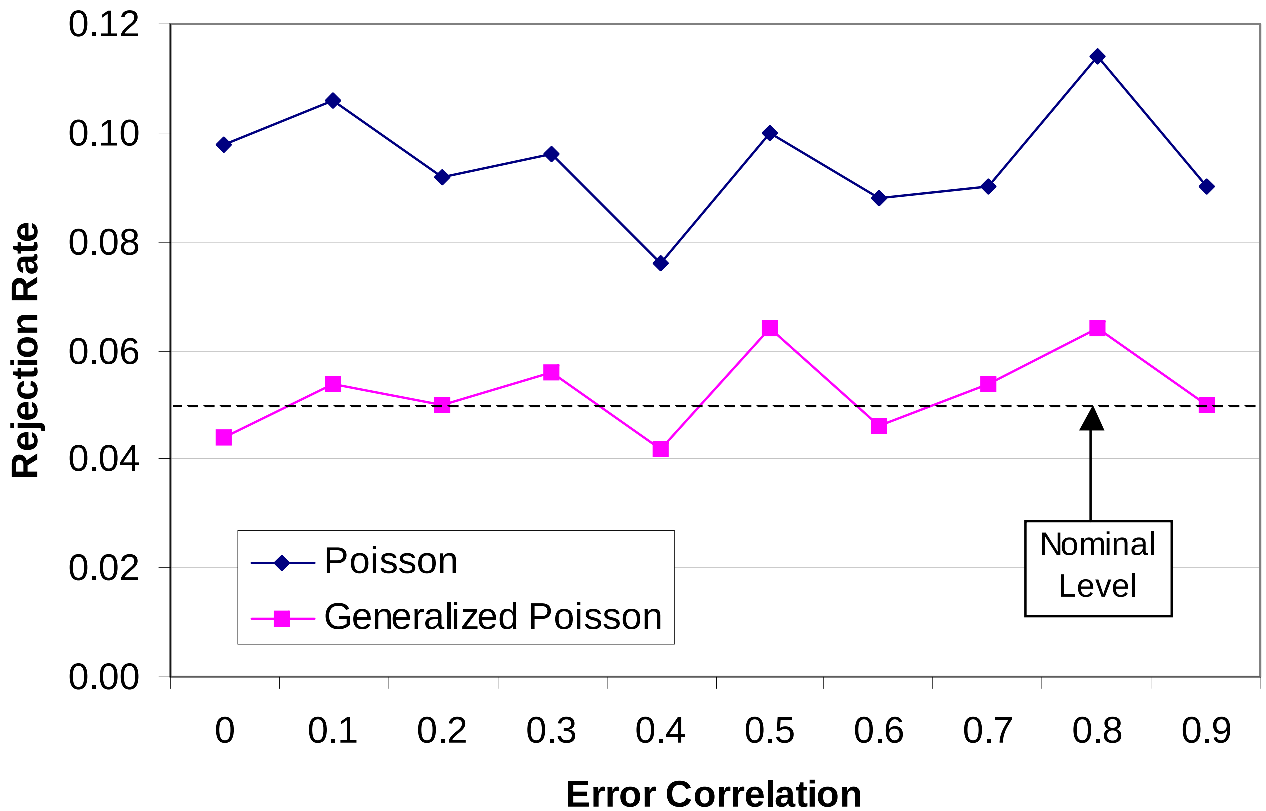
Simulation Results – I
ER Test for Unobserved Heterogeneity

- DGP $e_n = \sigma \times u_n$



Simulation Results – II
Wald Test for Lagrange Multiplier

- DGP $\beta_1 = 0$ with $e_n = 0.35 \times \underbrace{(\mathbf{I} - \rho_e \mathbf{W})^{-1} N(0,1)}_{\text{standardized}}$



OK! So, maybe it works with simulated data …

What about the real world … ?

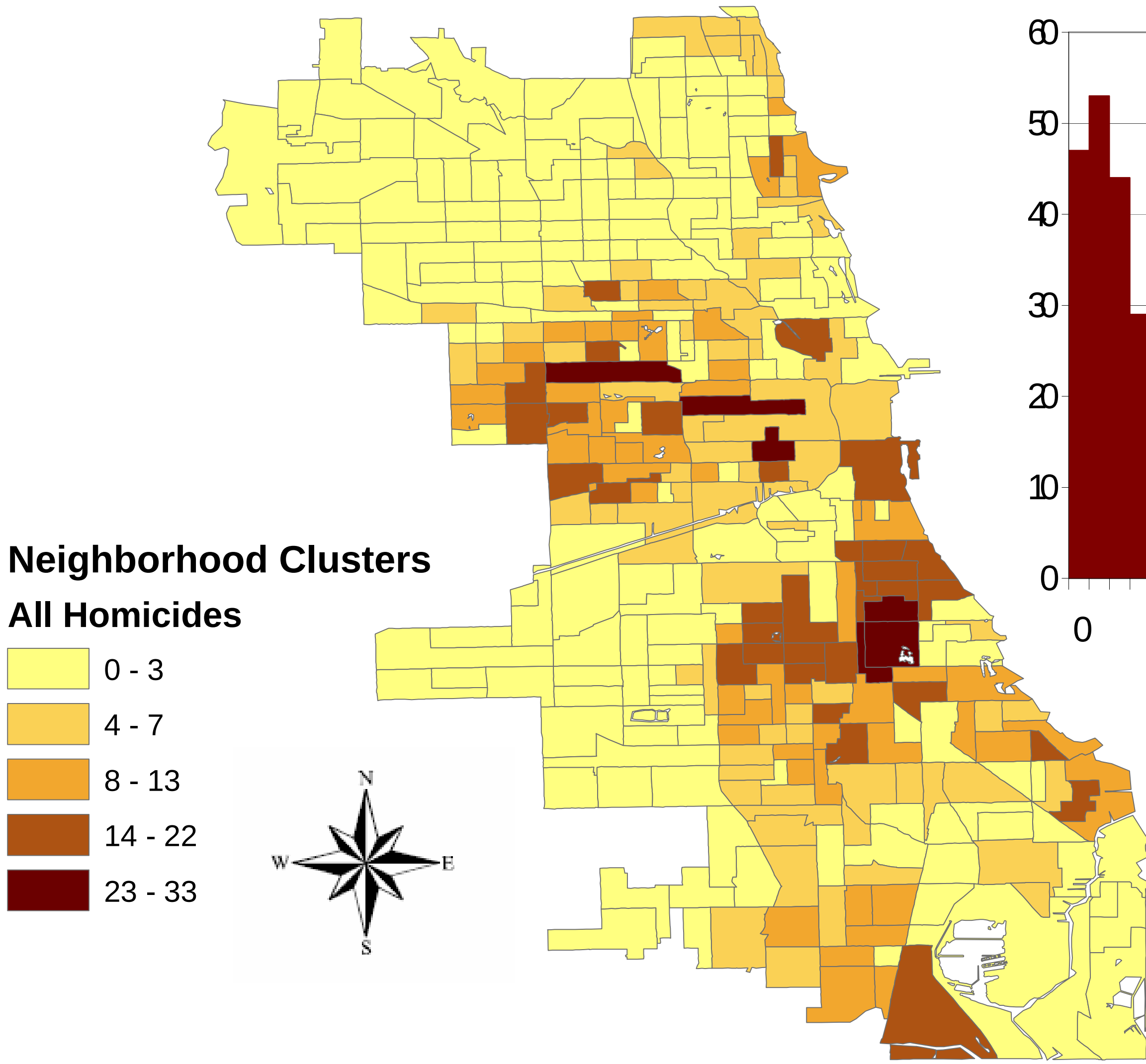
Application: Homicide Counts in Chicago Neighborhoods (1989-1990)

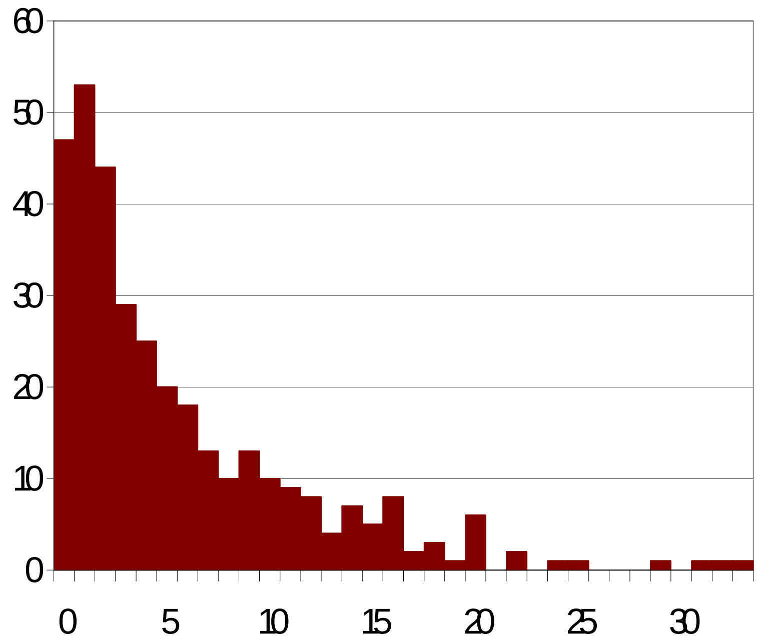
- Counts of homicides recorded in each of 343 Chicago neighborhood (PHDCN) between 1989-1991.

- Explanatory variables include (all census)
  - **LPOP**: Natural log of residential population (scale)
  - **RESDEP**: Resource depravation index (measuring concentrated disadvantage)
  - **RESST**: Percent of neighborhood households where the head of household has lived for more than 5 years (measuring residential stability)
  - **YMEN**: Young men as a % of population
  - **PNFH**: % of Non-family households

- Spatial Link Matrix: First Order Queen Criterion

Observed Spatial Distribution of the Homicide Count





Results:
All Homicides, Various Models

	Poisson	Negative Binomial	Zero-Inflated Generalized Poisson Model			Spatial Zero-Inflated Generalized Poisson Model		
			Amount of Violence	Predictability of Violence	Some Violence	Amount of Violence	Predictability of Violence	Some Violence
Intercept	-7.800 *	-8.380 *	-4.887 *	0.718	-7.373	-5.786 *	0.643	-1.424
LPOP	0.953 *	0.973 *	0.465 *	0.034	1.040	0.603 *	-0.012	0.342
RESDEP	0.714 *	0.816 *	0.841 *	-0.201 *	1.577 *	0.950 *	-0.244 *	0.725
RESST	0.151 *	0.198 *	0.127	-0.017	0.268	0.061	0.015	-0.044
YMEN	0.554 *	0.868 *	0.960 *	-0.341 *	-1.017	0.869 *	-0.319 *	-0.170
PNFH	0.124	0.238	0.384	-0.152	1.259	-0.156	0.093	2.389
Scale	…	0.171 *	…	…	…	…	…	…
ρ	…	…	…	…	…	0.211 *	0.467 *	-0.710
R^2	61%	59%			72%			74%

* = $p < 0.05$; + = $p < 0.1$

Results:
Expressive Homicides, Various Models

	Poisson	Negative Binomial	Zero-Inflated Generalized Poisson Model			Spatial Zero-Inflated Generalized Poisson Model		
			Amount of Violence	Predictability of Violence	Some Violence	Amount of Violence	Predictability of Violence	Some Violence
Intercept	-9.085 *	-9.298 *	-6.674 *	0.740	-1.390	-7.059 *	0.747	0.326
LPOP	1.076 *	1.057 *	0.609 *	0.055	0.467	0.661 *	0.035	0.278
RESDEP	0.825 *	0.938 *	0.928	-0.214 *	1.290 *	0.952 *	-0.227 *	1.283 *
RESST	0.028	0.059	-0.153	0.081	0.949	-0.147	0.081	0.956
YMEN	0.307 *	0.594 *	1.747 *	-0.816 *	-3.902 *	1.606 *	-0.760 *	-3.461 *
PNFH	0.192	0.355	0.838	-0.369	-0.732	0.867	-0.369	-1.042
Scale	…	0.178 *	…	…	…	…	…	…
ρ	…	…	…	…	…	0.084	0.218	-0.251
R^2	60%	58%			72%			73%

* = $p < 0.05$; + = $p < 0.1$

Results: Instrumental Homicides, Various Models

	Poisson	Negative Binomial	Zero-Inflated Generalized Poisson Model			Spatial Zero-Inflated Generalized Poisson Model		
			Amount of Violence	Predictability of Violence	Some Violence	Amount of Violence	Predictability of Violence	Some Violence
Intercept	-8.691 *	-8.506 *	-4.122	-2.271	-0.885	-3.372	-3.779 *	-1.077
LPOP	0.902 *	0.862 *	0.429	0.294	-0.058	0.417 +	0.406 *	-0.031
RESDEP	0.787 *	0.855 *	0.749 *	-0.179	0.237	0.260 *	0.263 *	0.794 *
RESST	-0.133	-0.147	-0.084	0.048	-0.307	0.239	-0.187	-0.687 +
YMEN	0.697 *	0.865 *	0.356	-0.120	1.271	0.100 +	0.215	1.510 +
PNFH	0.349	0.468	-0.573	0.418	1.941	-0.642 *	0.790 *	1.677
Scale	…	0.243 *	…	…	…	…	…	…
ρ	…	…	…	…	…	-0.782 *	0.955 *	0.320
R^2	42%	42%			45%			51%

* = p < 0.05; + = p < 0.1

Interesting Model Implications

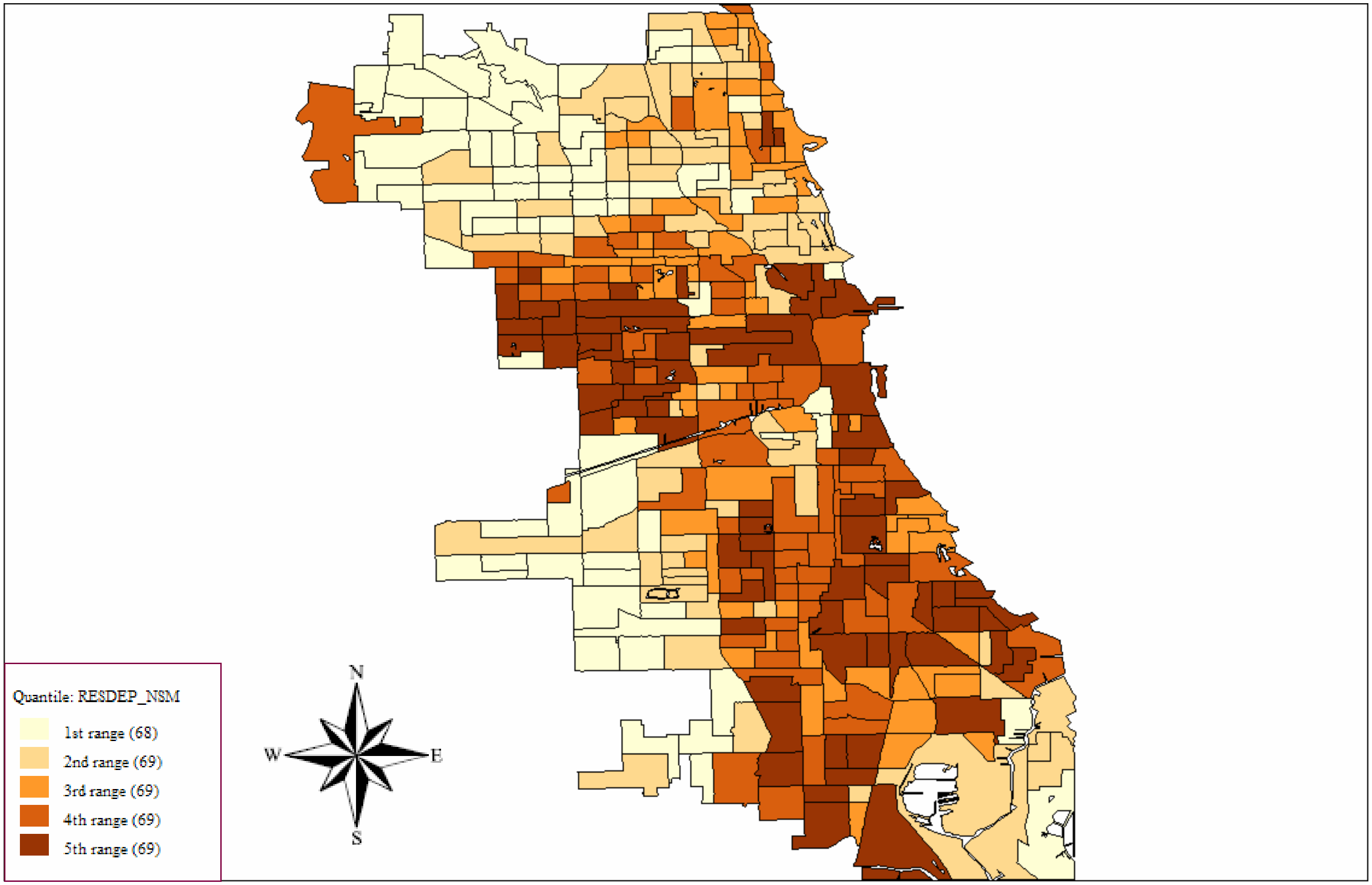
- Tease out decomposition of marginal effects

$$\frac{\partial \ln(\hat{s}_n)}{\partial x_{kn}} = \frac{\hat{\phi}_n(1,1)}{\hat{s}_n}\lambda_{k1} + \frac{\hat{\phi}_n(1,2)}{\hat{s}_n}\lambda_{k2} + \frac{\hat{\phi}_n(1,3)}{\hat{s}_n}\lambda_{k3}$$

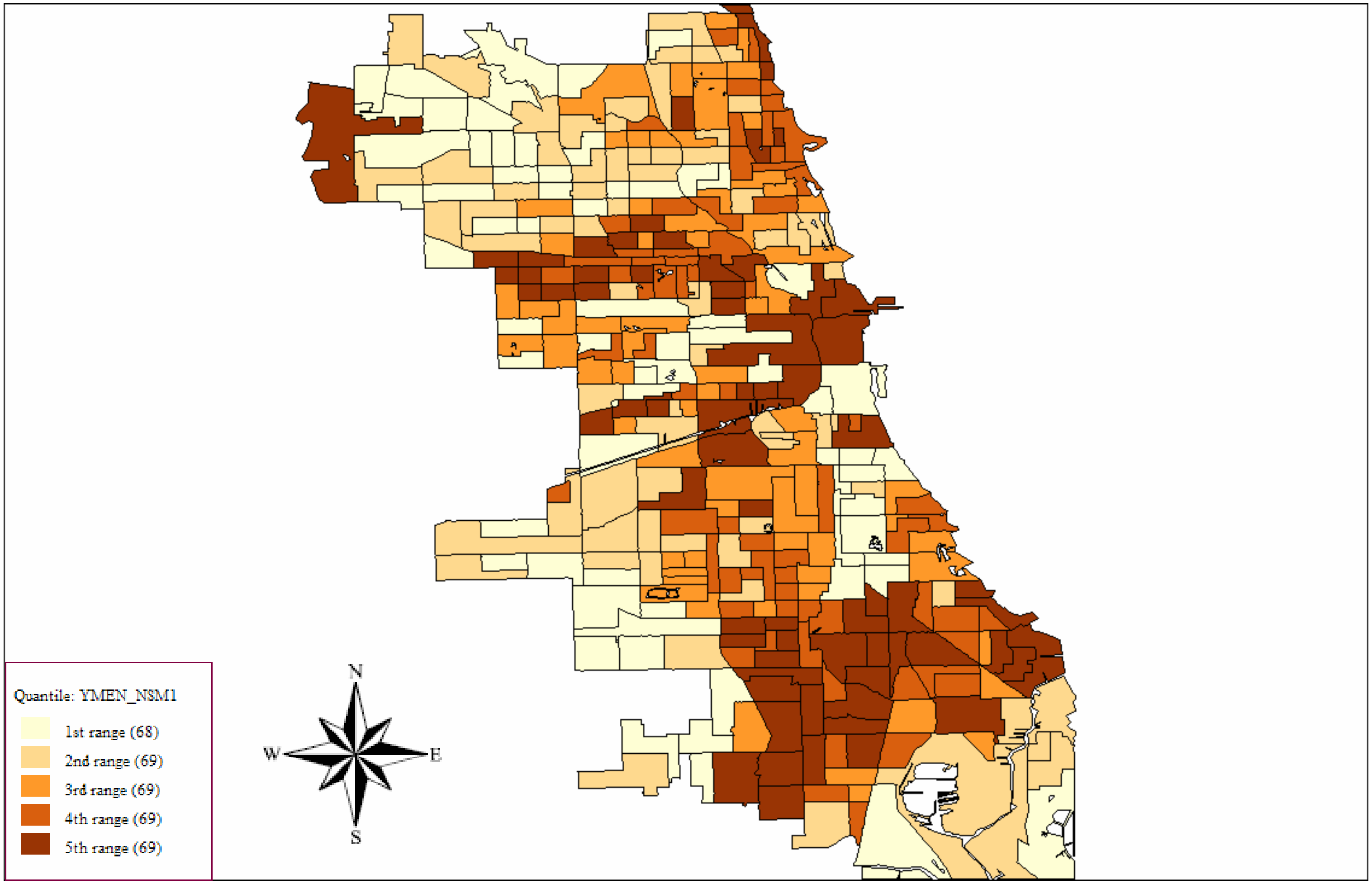
  where $\hat{\phi}(i,j)$ is the covariance between the i^{th} and j^{th} moments.

- Examine variation in marginal effects across space
  - Map out variations in effects
  - Study what correlates with high/low effects

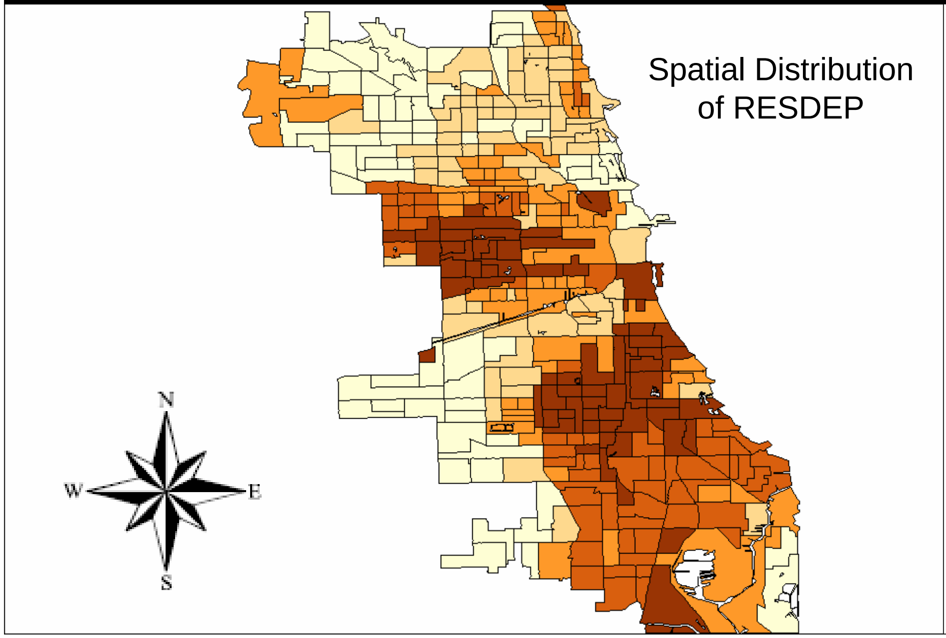
Spatial Variation in the Effect of RESDEP

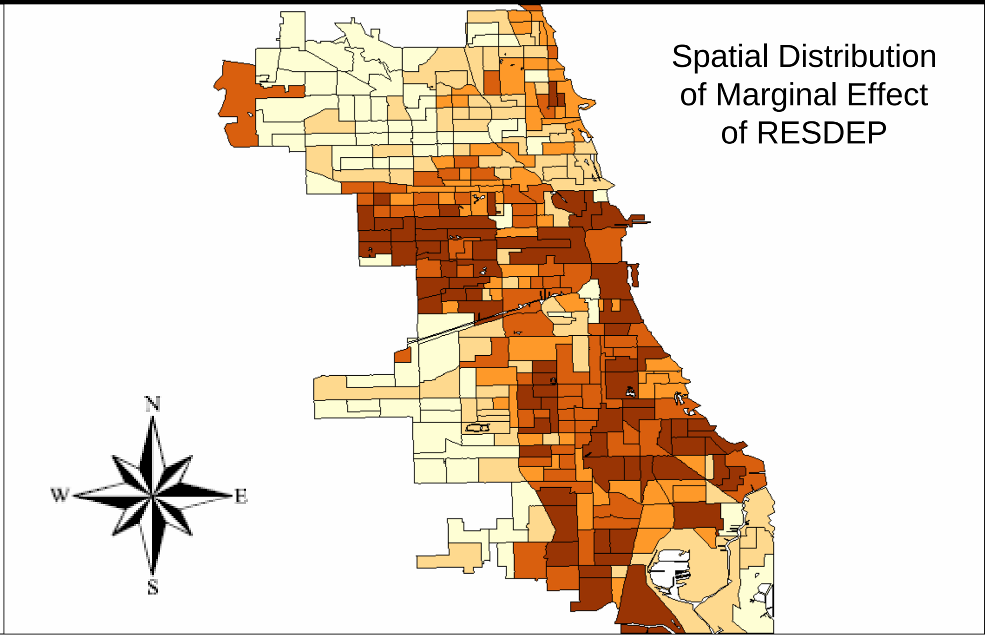


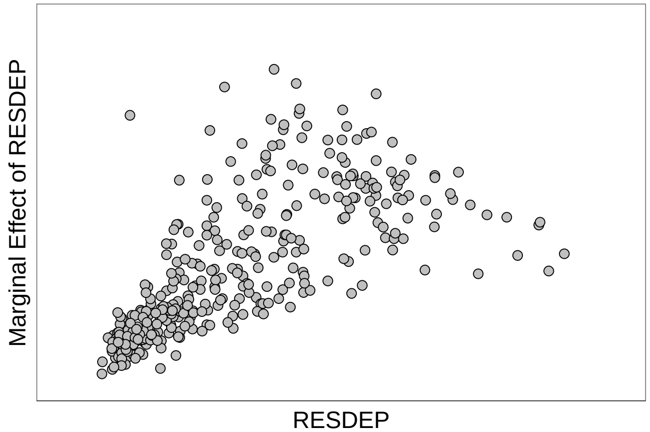
Spatial Variation in the Effect of YMEN



Marginal Effects of RESDEP versus RESDEP

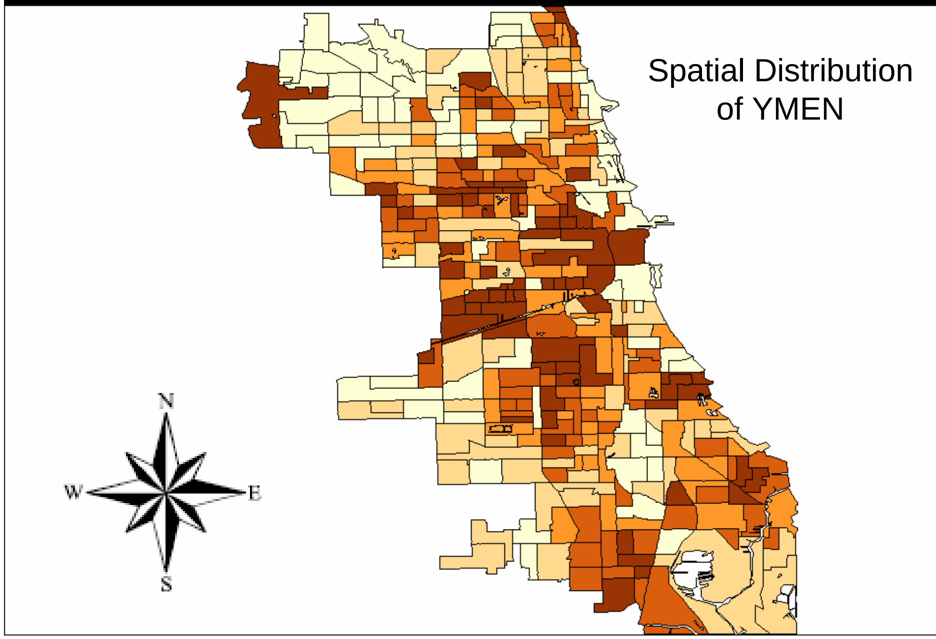


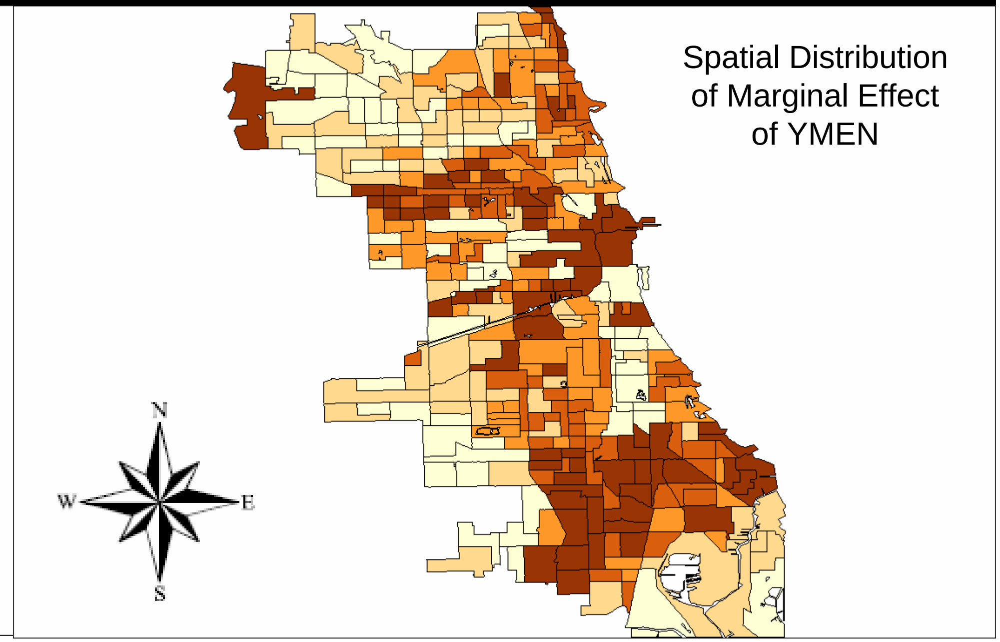


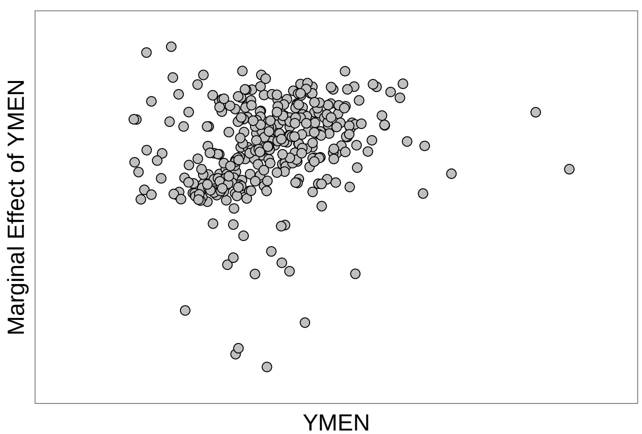


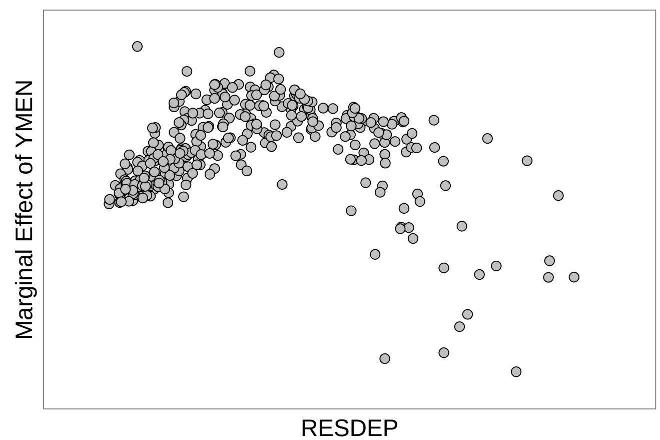
URBAN INSTITUTE
Justice Policy Center

Marginal Effects of YMEN versus YMEN and RESDEP



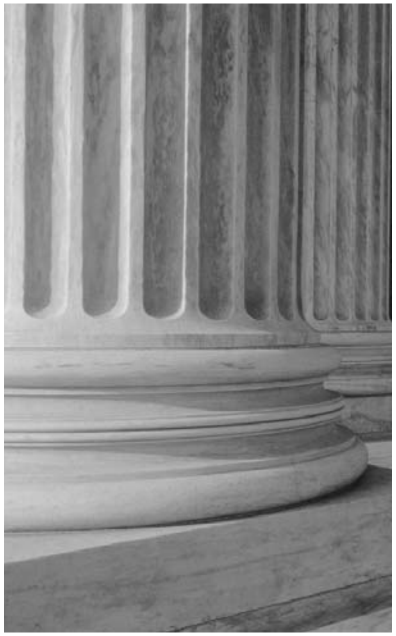






Concluding Thoughts

- Approach presented here has several good features
  - Minimal distributional assumptions invoked a-priori
  - Extends easily to incorporate several real-word data/sample features
  - Other non-sample knowledge can be included
  - Yields varying effects across space
  - Can yield very precise policy recommendations
- Further Extensions
  - Endogenous predictors (e.g., resource depravation ?)
  - Simultaneous Equation Count Models (e.g., different types of homicides)
  - Space-time models
- To do …
  - More Simulations
  - More cross method comparisons (e.g., Bayesian)



URBAN INSTITUTE
Justice Policy Center

Acknowledgements:

- **<u>Organizational support</u>, Justice Policy Center, The Urban Institute**
- **<u>Moral Support and Patience</u>: Wife and Kids**

Contact Information:

Avinash Singh Bhati
(202) 261-5329
abhati@ui.urban.org