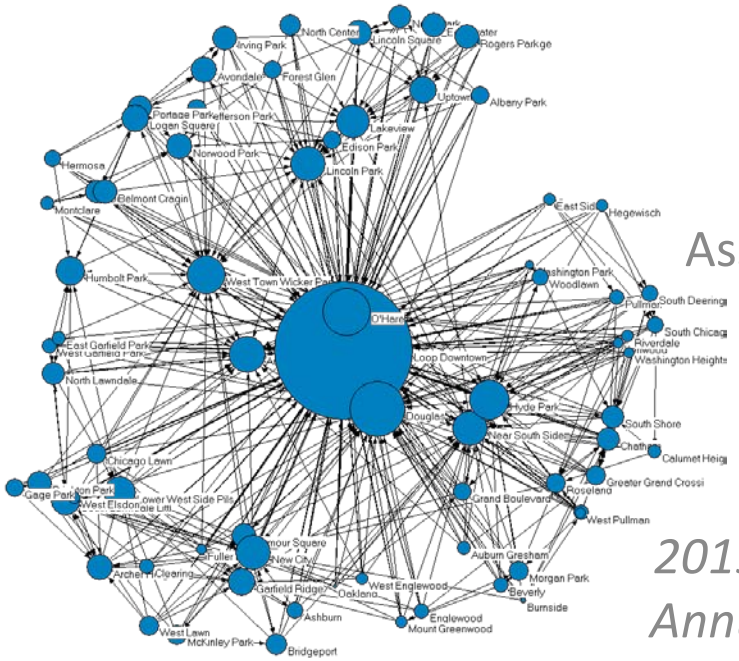
Spatial Network Links to Maternal and Child's Health: New Insights from Linking Longitudinal Employer-Household Dynamics (LEHD / LODES) Data with Vital Record Statistics



Corina Graif,

Associate Professor of Sociology and Criminology

Collaborators: Megan Evans and Stephen A. Matthews, Pennsylvania State University

Presentation at the:

2019 Local Employment Dynamics (LED) Partnership Annual Workshop, Sept 4-5, 2019, US Department of Commerce and US Census Bureau

Acknowledgements

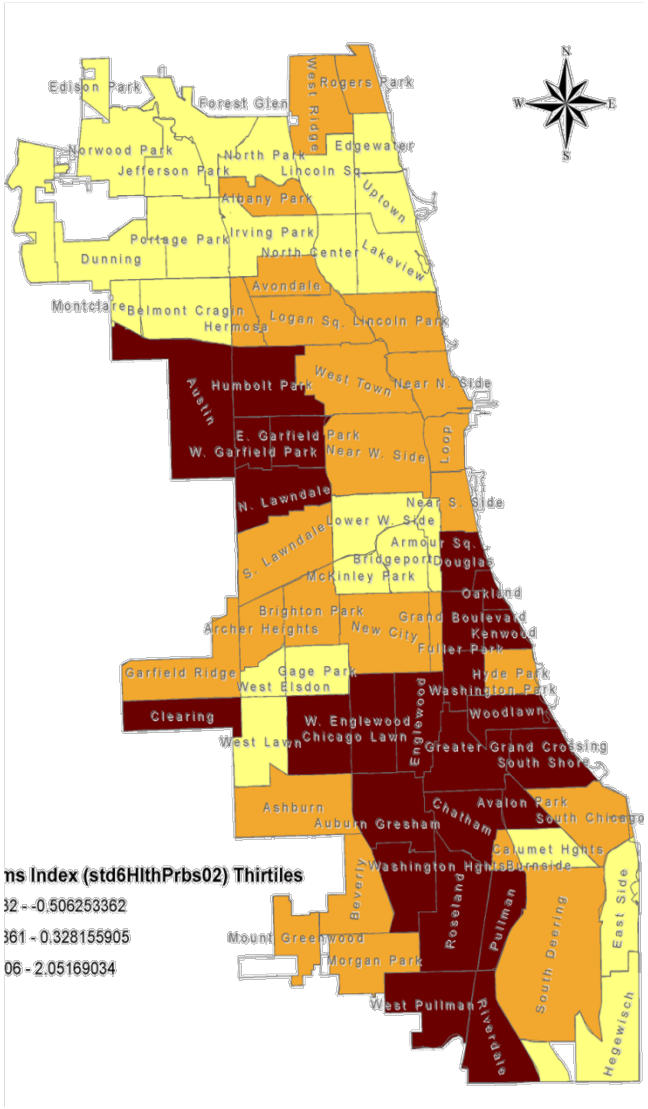
Thank you for funding to:

- National Institute of Health (NICHD K01-HD093863)
- Penn State University's Population Research Institute (NICHD R24-HD041025)
- National Science Foundation (NSF CNS-1544455)

Neighborhood Effects on Health Risk

- **Background:**
  - Neighborhoods affect many health risk behaviors and outcomes

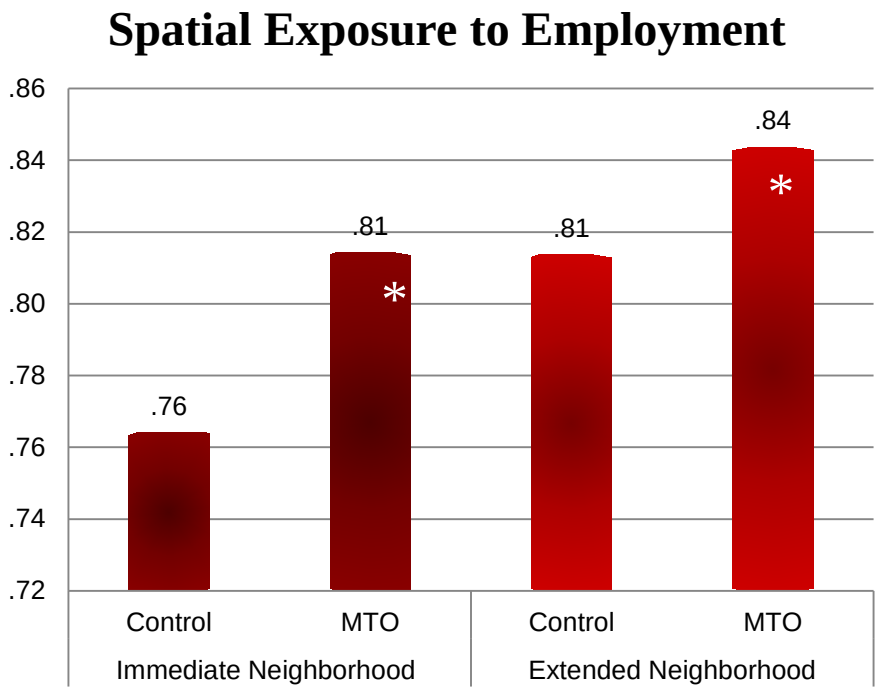
    E.g., Neighborhood poverty affects
    - aggression, cognitive functioning, physical health,
    - mental health (Graif et al 2016),
    - victimization (Graif and Matthews 2017),
    - risk-taking (Graif 2015)
- **Puzzle:**
  - Neighborhood interventions, like MTO, unexpected findings
    - E.g. moving to affluence increased risk taking of boys (even as it decreased that of girls)

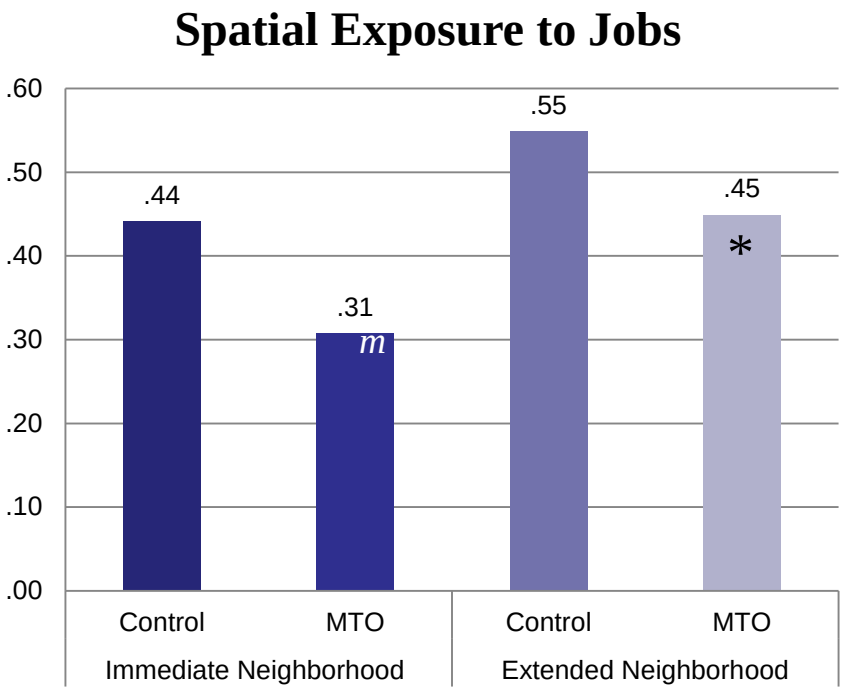


Possible core source of the puzzle:

- Neighborhoods assumed to be **independent** of each other
  - predominant working assumption in many studies
  - problematic (e.g., spatial interactions, activity space research, diffusion research)
- Unmeasured variables:
  - Respondents' differential exposures to **non-residential places** during daily activities
  - but only *residential places* measured
- But collecting neighborhood connectivity and daily mobility data is expensive

Link different data sets e.g. LEHD, CENSUS, and MTO





Paradox: decreases in unemployment do not translate into increases in spatially proximate job presence.

The immediate and extended job environment is significantly worse for the MTO group

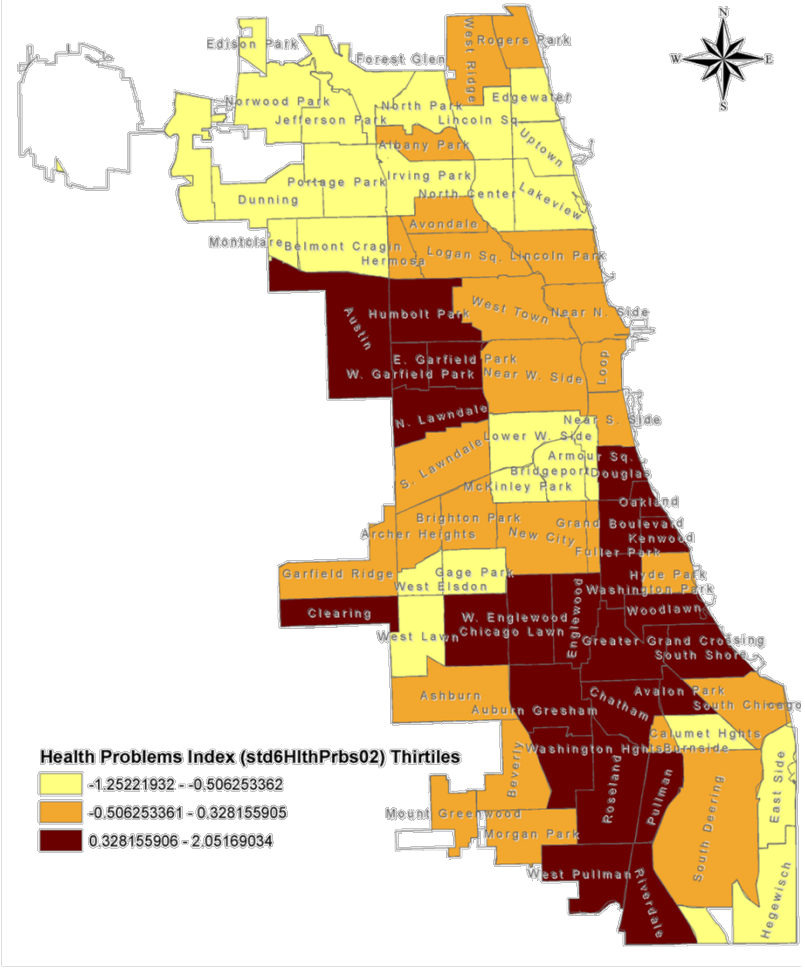
Major implication: What is a "good" neighborhood?

Ref: Graif, C. 2011. *Mobility in Isolation*. Harvard University

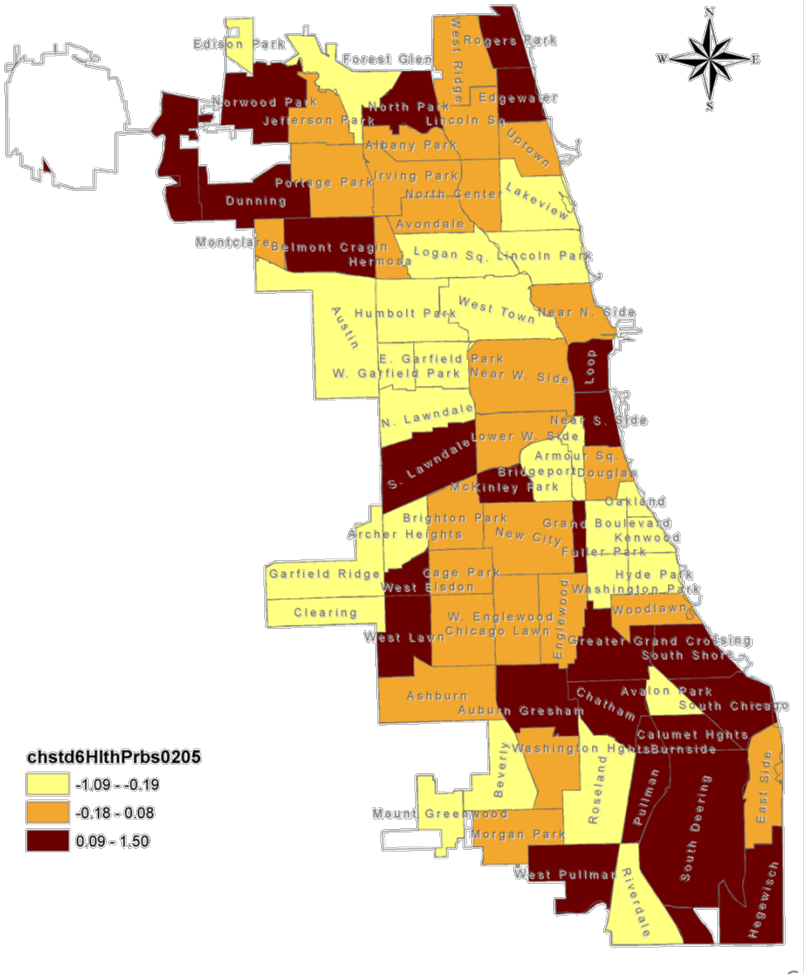
Spatial Clustering of Health Problems

(infant mortality, low birth-weight, very low birth weight, births to teen mothers, lead poisonings)

Health Problems 2002



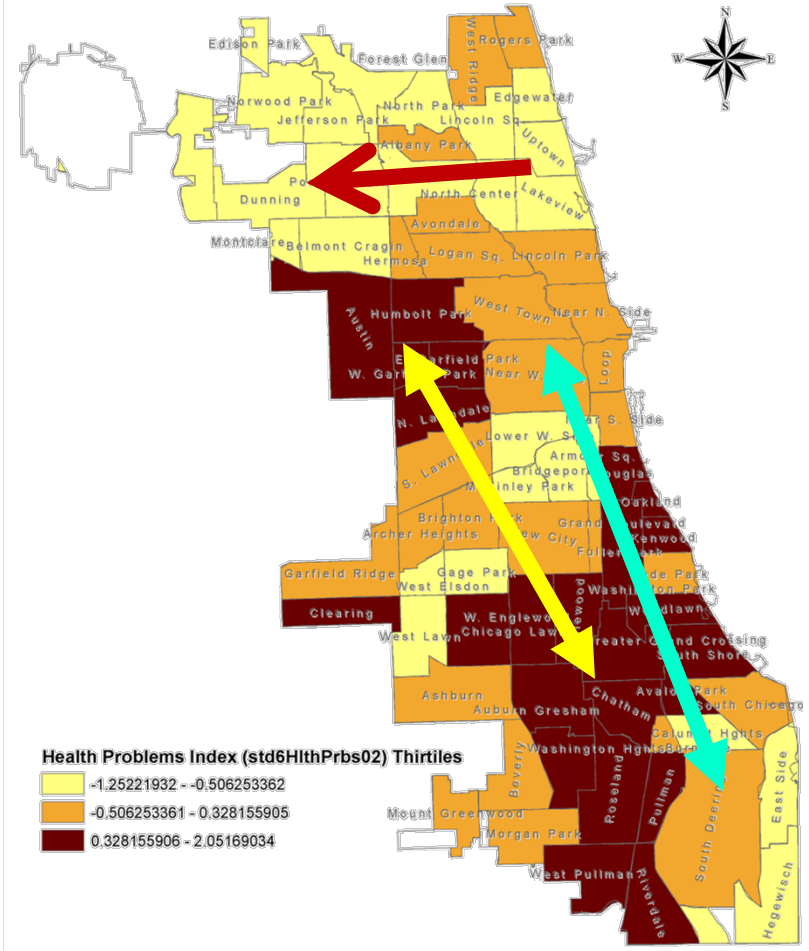
Changes 2002-2005



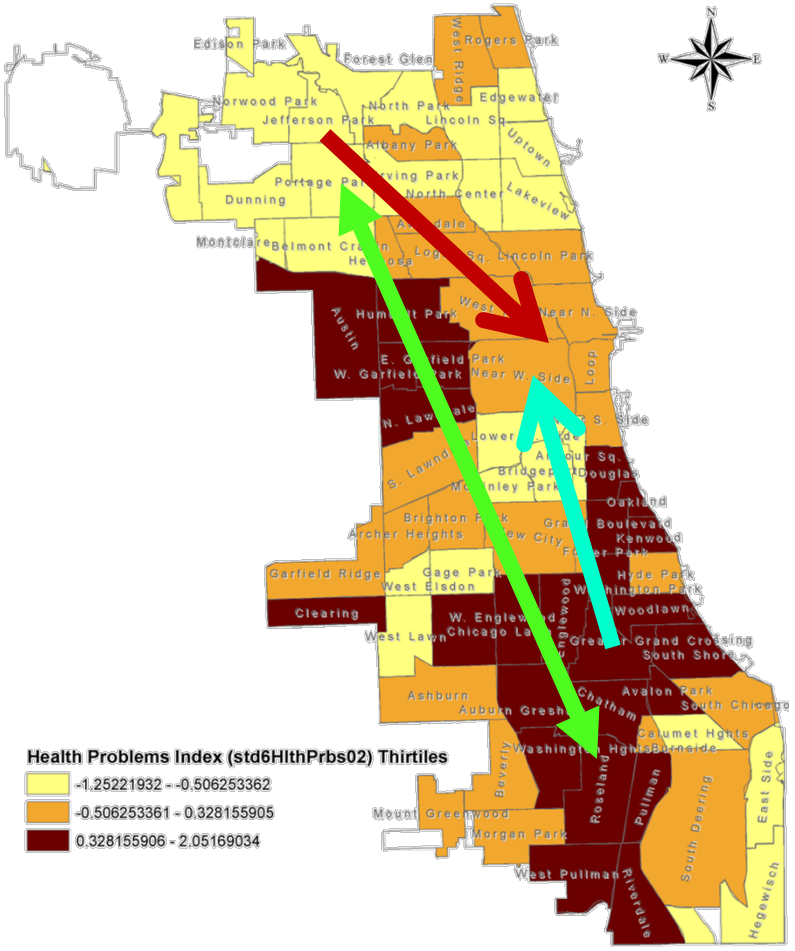
Research Question:

Are commuting ties associated with neighborhood maternal and child health problems?

Health Problems 2002



Network connections



Measures and Data

- Neighborhoods:
  - Chicago Community Areas (N=77), aggregations of Census tracts, Census Tiger shape files
- Neighborhood outcomes
  - Vital record statistics: infant mortality, pre-term births, low birth weight, and very low birth weight
    - (also, lead poisonings, teen births, violent deaths)
  - Police reports of crime incidents 2001 - 2013
    - Violent crime: homicide, robbery, assault, battery, sexual assault, domestic violence
    - Property crime: burglary, crim damage, theft, motor vehicle theft

Measures and Data

- Neighborhood socioeconomic and institutional characteristics
  - Decennial Census and American Community Survey
    - Poverty, unemployment, public assistance, female headed households (factor score): concentrated disadvantage index
    - Residential stability (home ownership, residents who lived in the area for 1-5 years)
    - Racial and ethnic composition, racial and ethnic diversity index (herfindahl concentration score)
  - Land use measures (LEHD)
  - Transportation network (Chicago Data portal)

Network and Spatial Measures and Data

- Inter-neighborhood connections based on commuting flows:
  - Longitudinal Household Employment Dynamics 2001 - 2013 (On the map),
  - OD, WAC, RAC
  - Ties are weighted by the proportion of all commuters in home CA
- Dyadic measures:
  - Geographic proximity matrix (using Tiger shape files)
  - Network proximity matrix
  - Inter-neighborhood similarities in disadvantage, and race/ethnicity, health problems,

Methodology

- Spatial analysis
  - Exploratory mapping; Spatial proximity matrix;
  - GIS analysis
  - Examine job hubs and institutional and job deserts

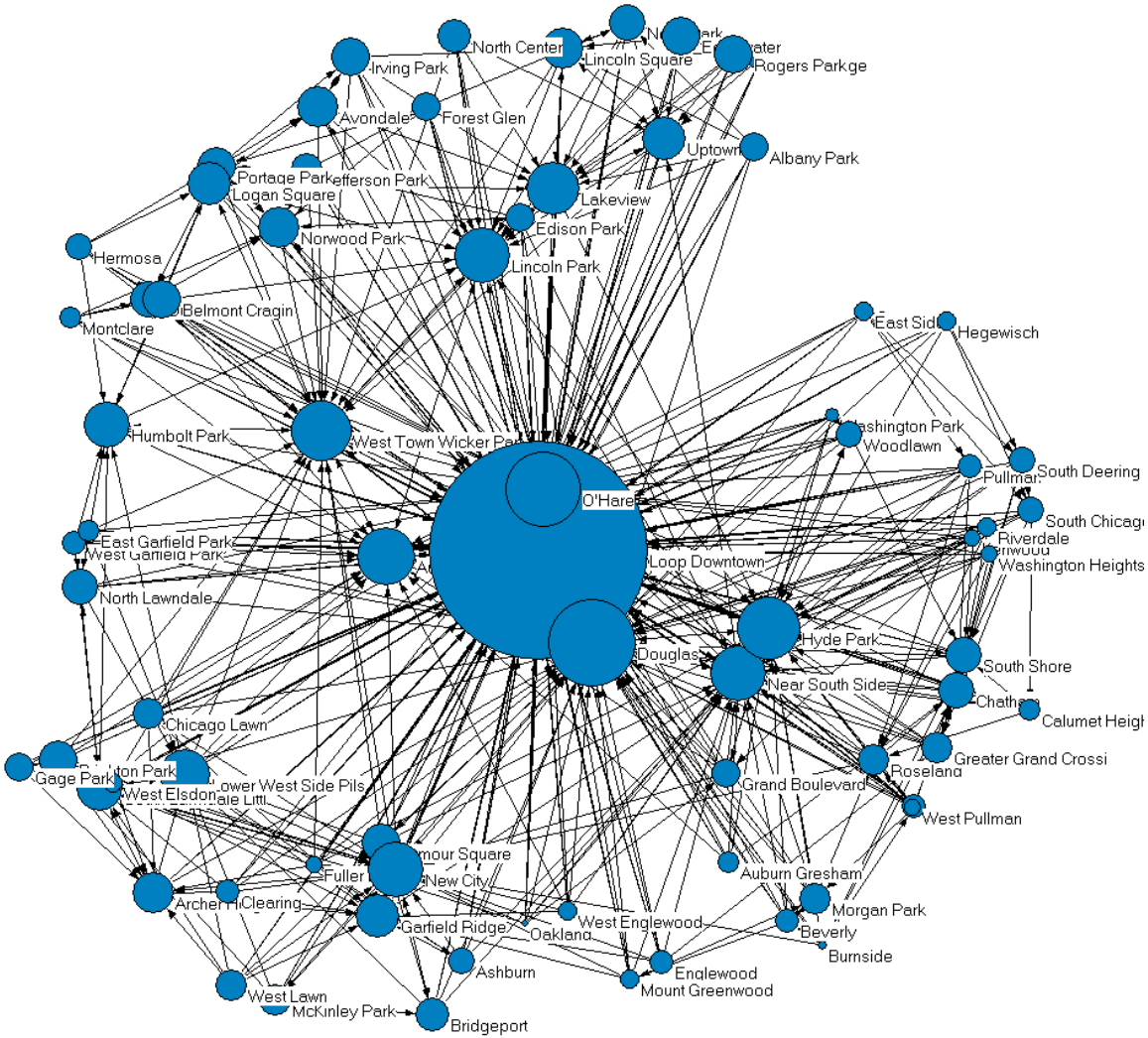
- Network analysis
  - Created inter-neighborhood networks based on commuting flows (overall and by different characteristics of the commuters)
  - Combined with GIS analysis and mapping
  - Network weighting of inter-neighborhood dependencies based on commuting ties;

- The analytical unit is defined at the **dyadic level**
  - (the ties between two community areas).
  - LEHD Origin Destination data used to define the *network ties*
  - Census, ASC, RAC and WAC data for attributes of the community areas = *network nodes*

Methodology

- Multiple Regression Quadratic Assignment Procedure (MRQAP in UCINET),
- ERGM (exponential-family random graph models), and
- TERGM (temporal ERGM) using R statnet
- Negative binomial regression of counts
- Leave-one-out classification methods
- Permutations

Network Configuration of Chicago Communities



Legend

Circle size = CA in-degree

Line = commuting tie (dychotmized here but weighted in analyses)

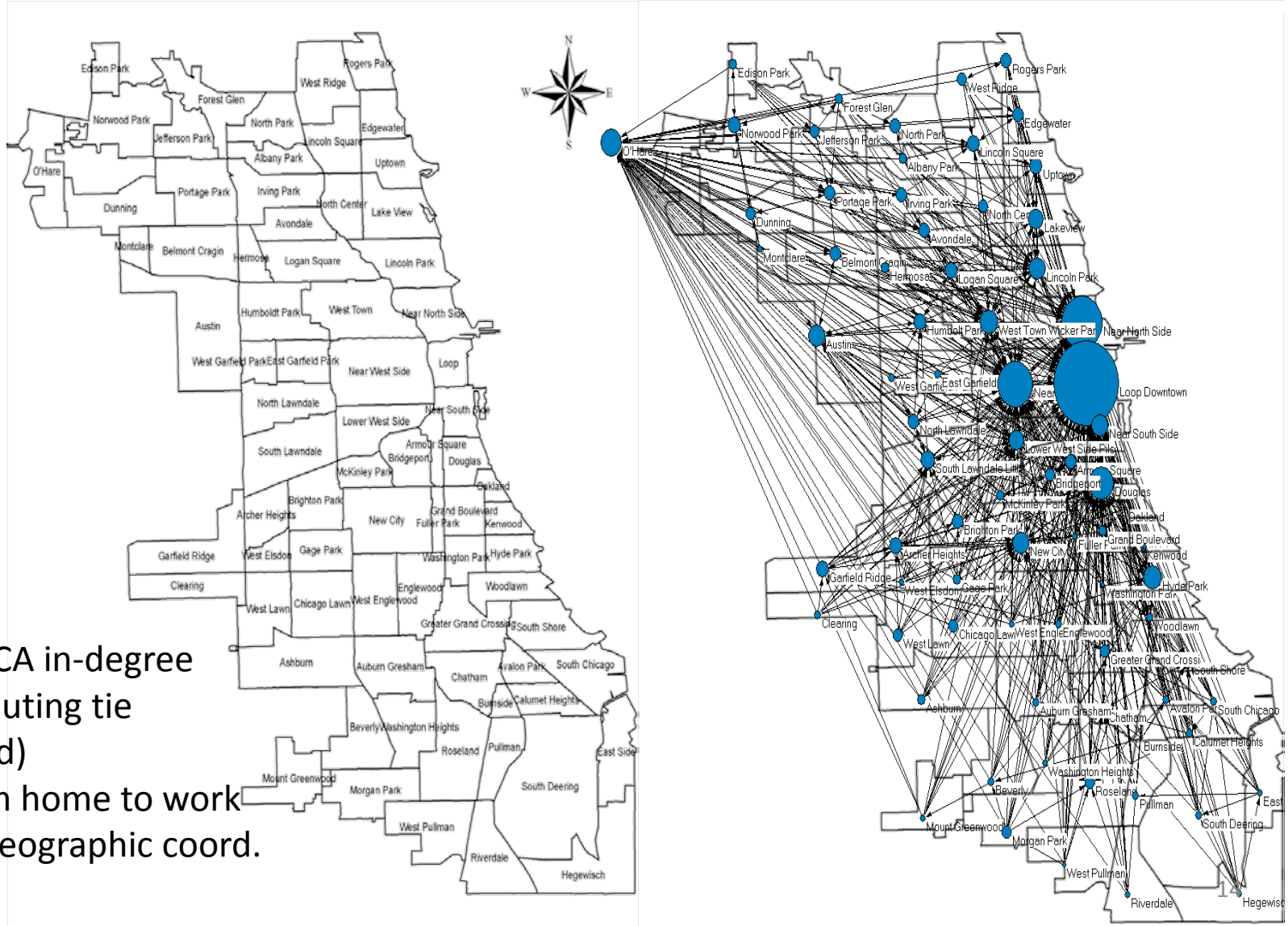
Arrow = from home to work

Location = Fruchterman Reingold F1 (non-geographic)

Note: Showing only up to 10 highest value arcs per home CA

Source: Author's network analysis of 2002 LEHD data aggregated to CA-by-CA level

Integrating Geography and Networks



Legend:

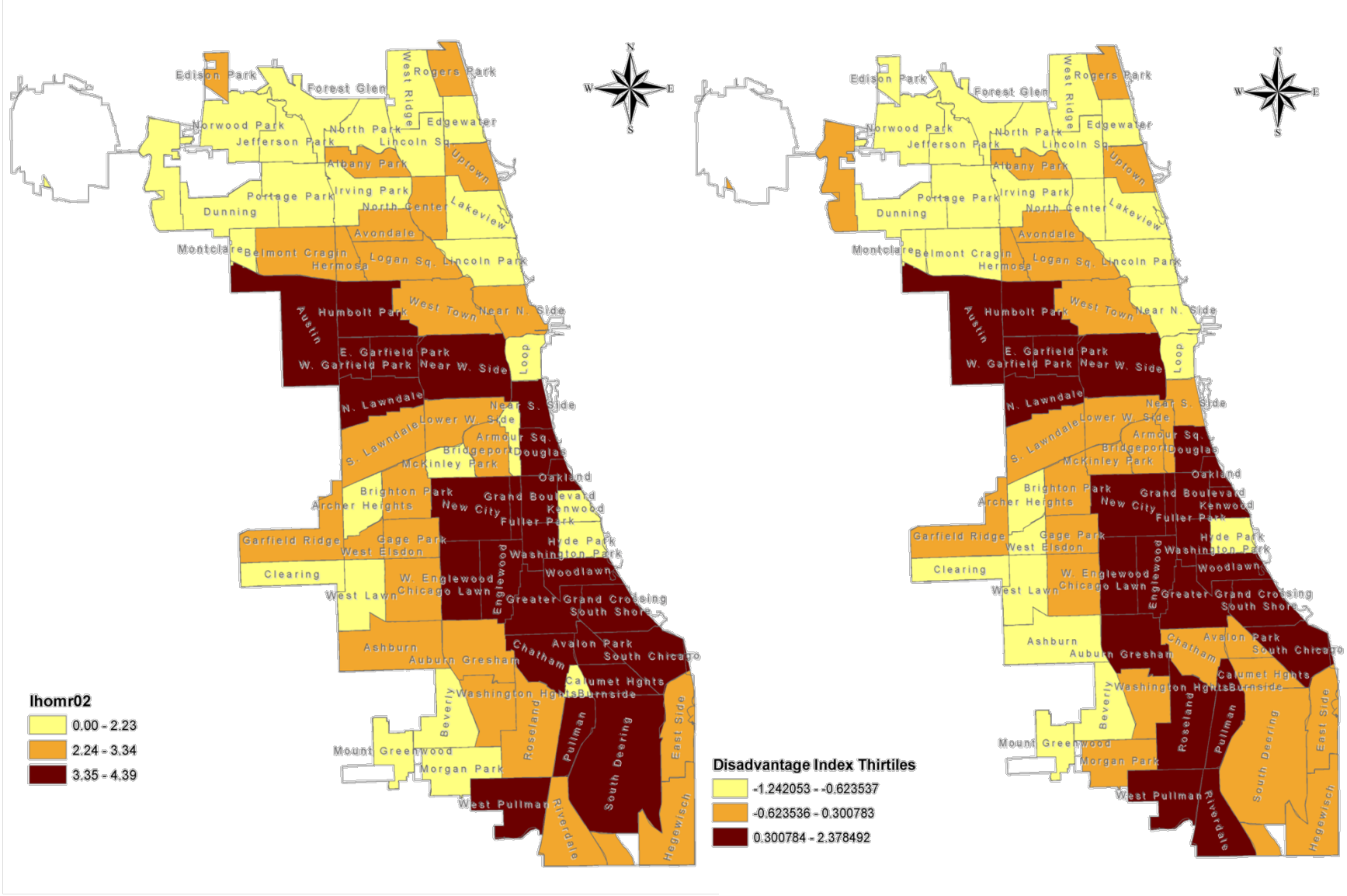
Circle size = CA in-degree

Line = commuting tie (dychotmized)

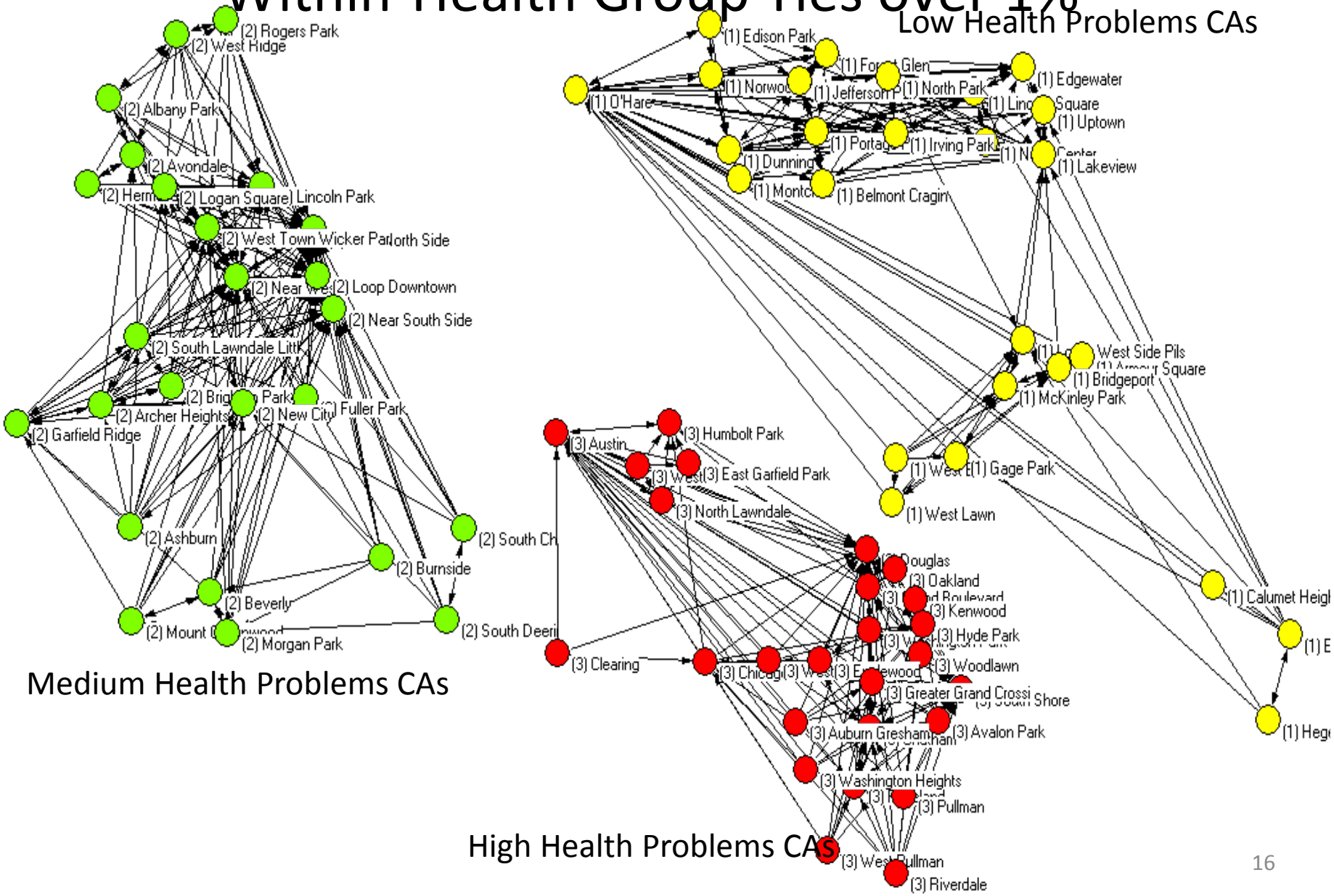
Arrow = from home to work

Location = Geographic coord.

The Geography of Violent Crime



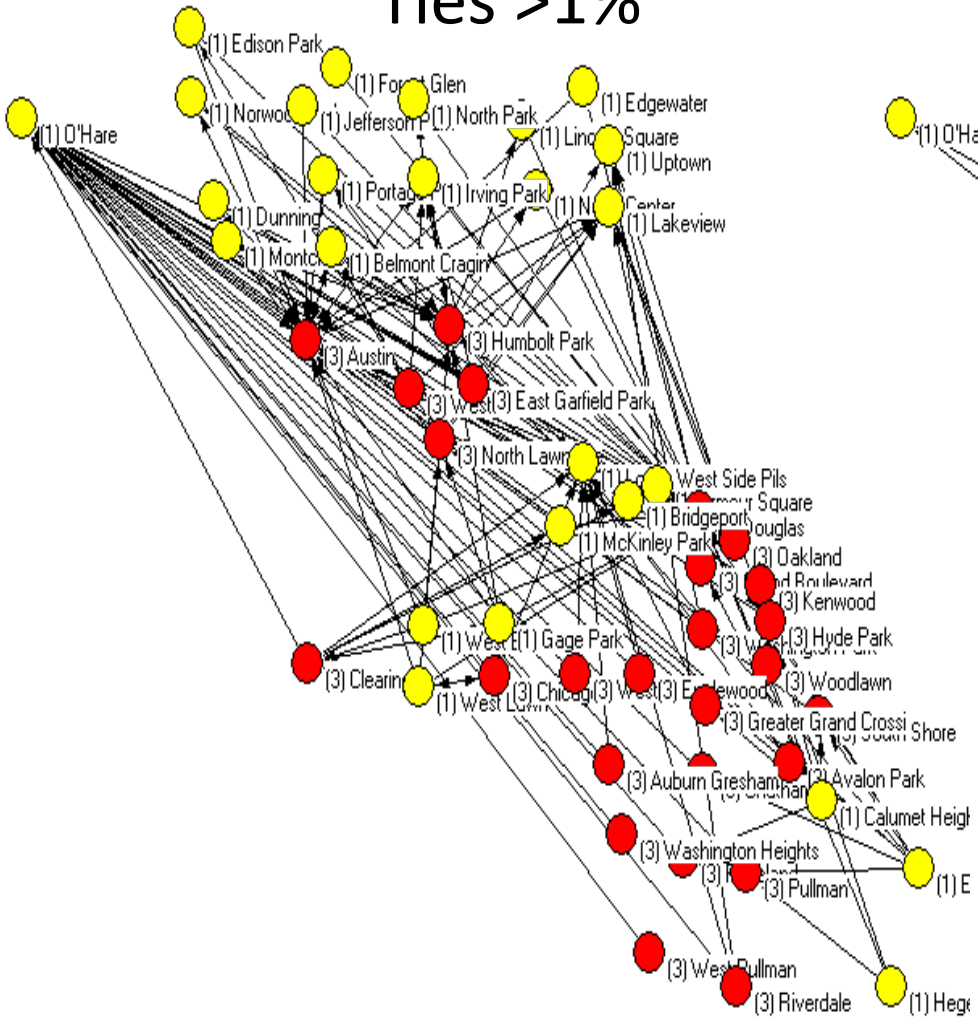
Bad News: Strong Homophily Within-Health Group Ties over 1%



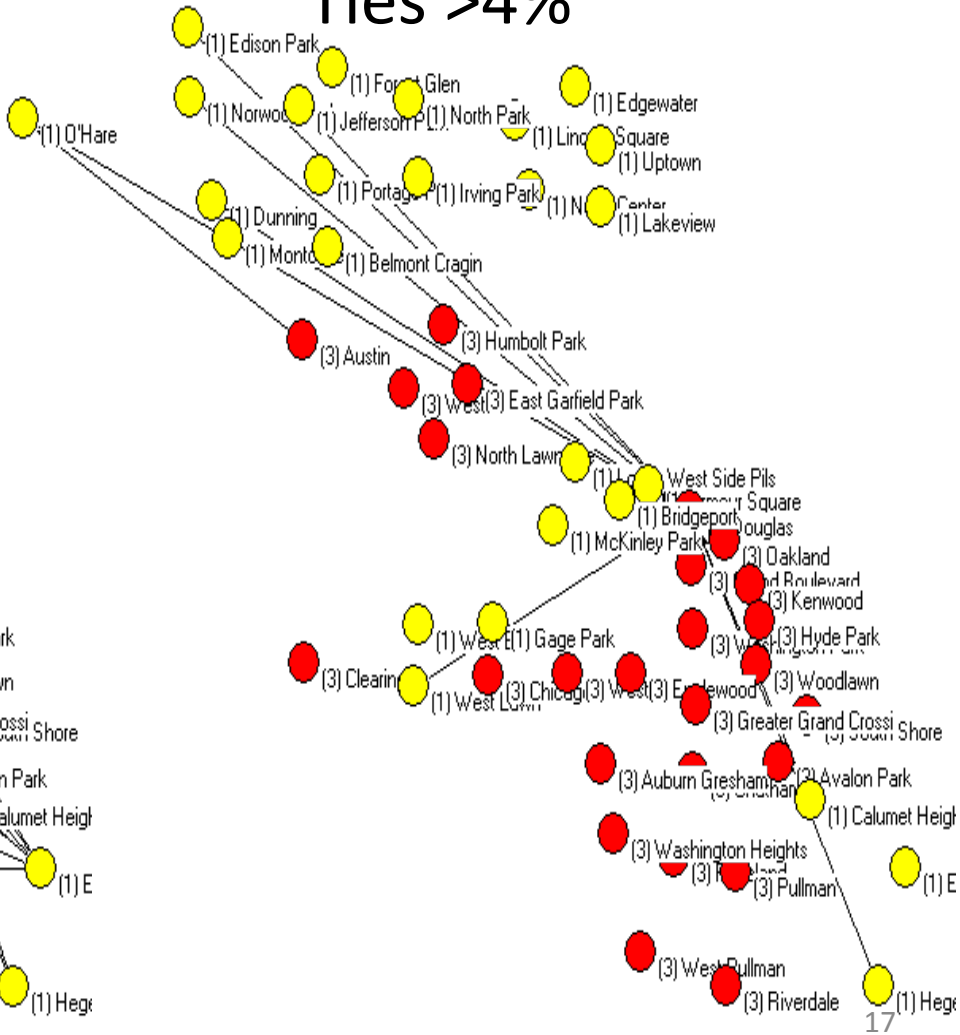
Good news: Heterophily Between Health Group Ties

Low-Health Problems CAs by High-Health Problems CAs

Ties >1%



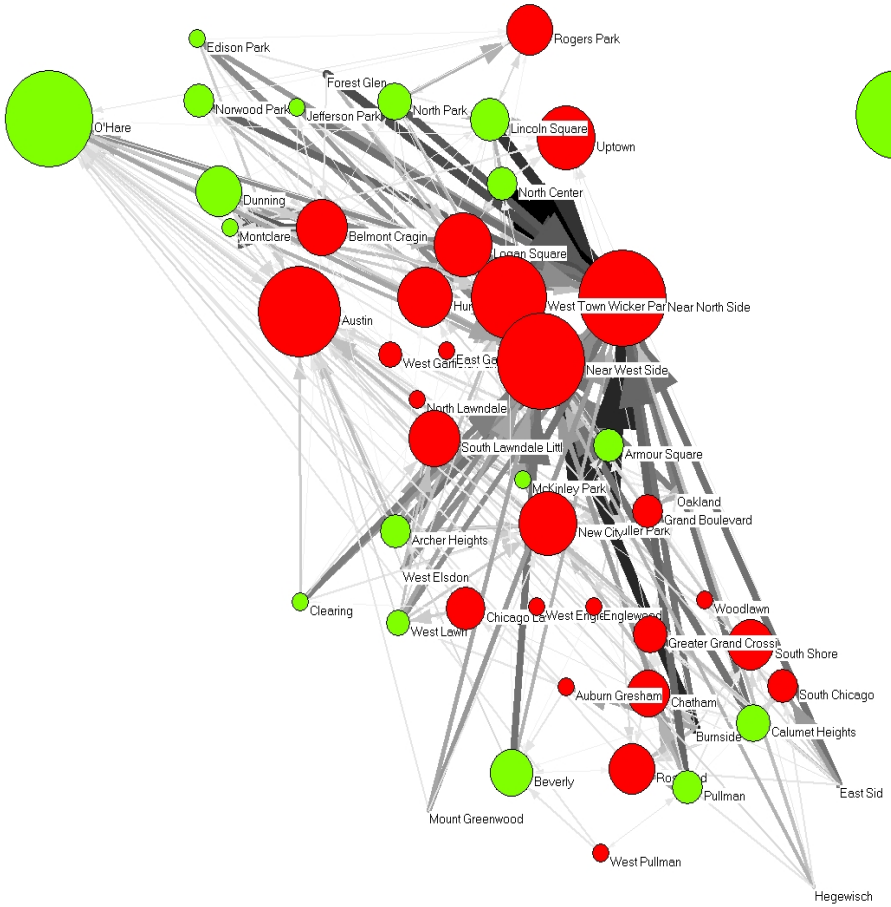
Ties >4%



Good news???: Network Heterophily

inter-neighborhood ties (low income commuters)

--only ties between *high (red) and low (green) neighborhood violence ;* **between-group** ties (>1%, 2%-4%); vertex size -- indegree



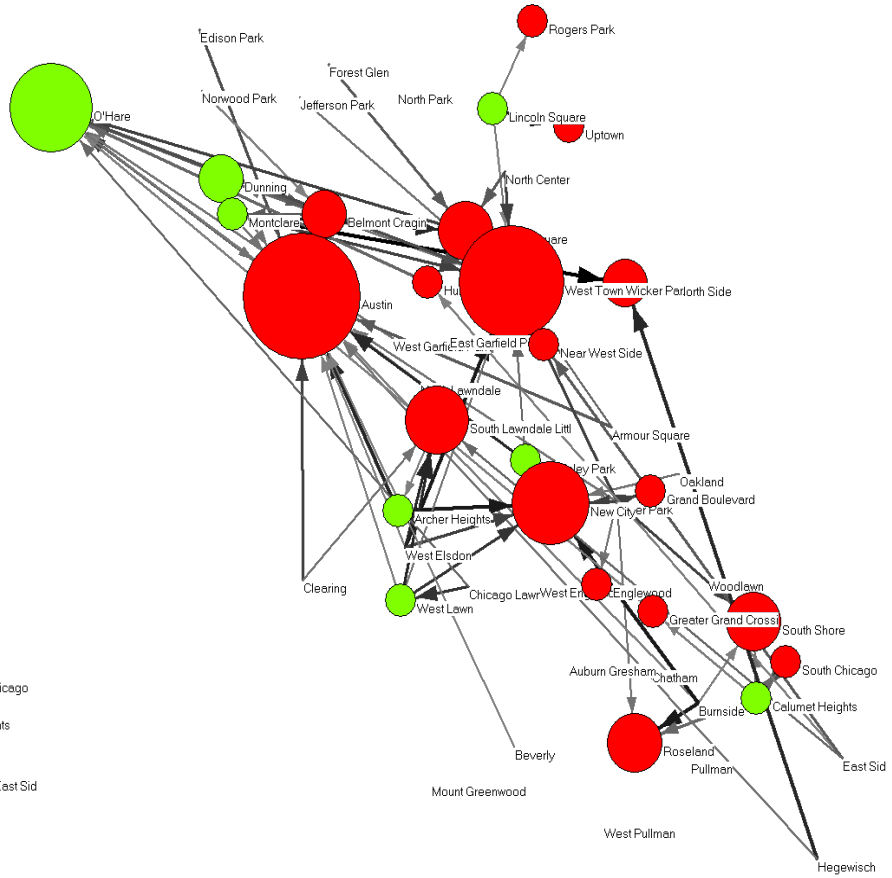


Table 1: Multiple Regression (Quadratic Assignment Procedure) Estimating Pairwise Similarity in Health Problems

	M1	M2	M3	M4	M5
Directed network ties*	.098 ***	.078 *	.064 *	.051 *	.056 *
	(.001)	(.015)	(.025)	(.047)	(.033)
Similarity in disadvantage		.591 ***	.584 ***	.340 ***	.320 ***
		(.001)	(.001)	(.001)	(.001)
Geographic proximity			.090 **	.009	.001
			(.010)	(.379)	(.469)
Similarity in % non-Hisp Black				.404 ***	.430 ***
				(.001)	(.001)
Similarity in % Hispanic					-.068 *
					(.037)
AdjRSq	.010 ***	.359 ***	0.367 ***	0.459 ***	0.463 ***
p-val	(.001)	(.001)	(.001)	(.001)	(.001)
N ties	5852				

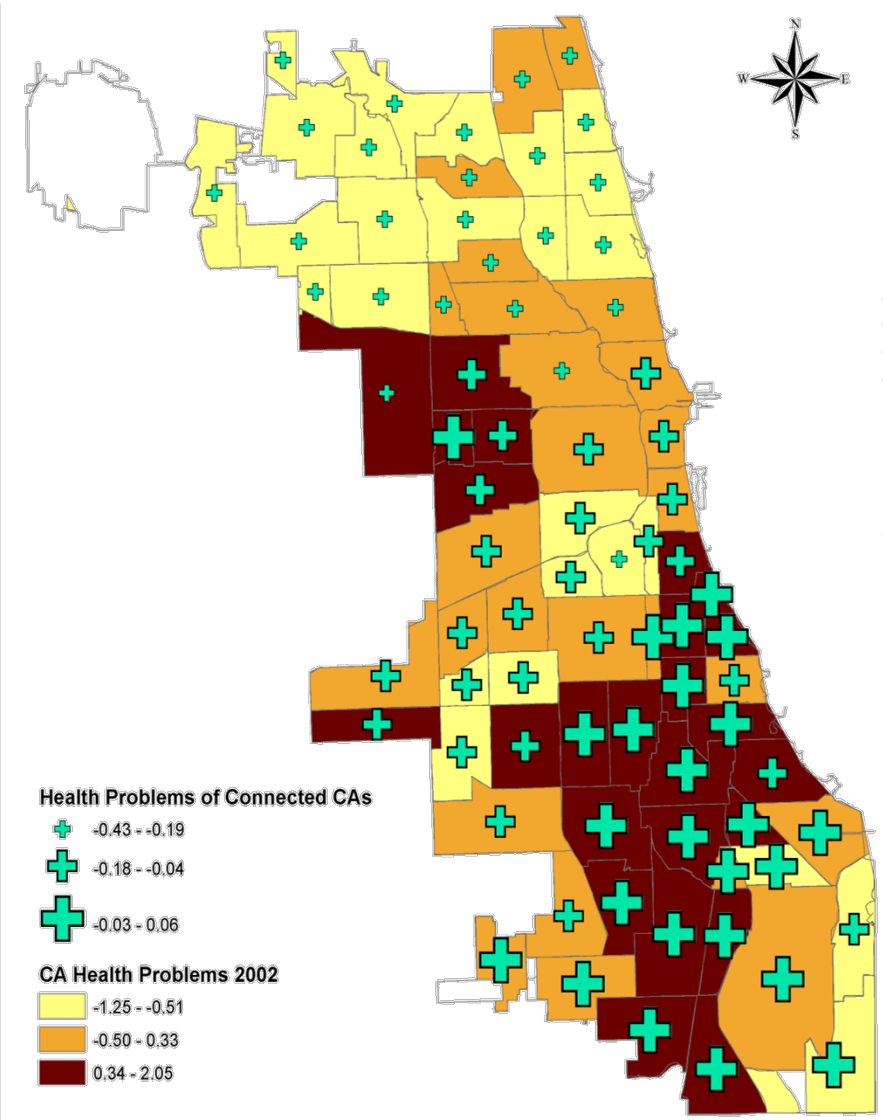
Note: Cells represent standardized coeffcients and p-value in parantheses; estimations are based on 2000 permutations;

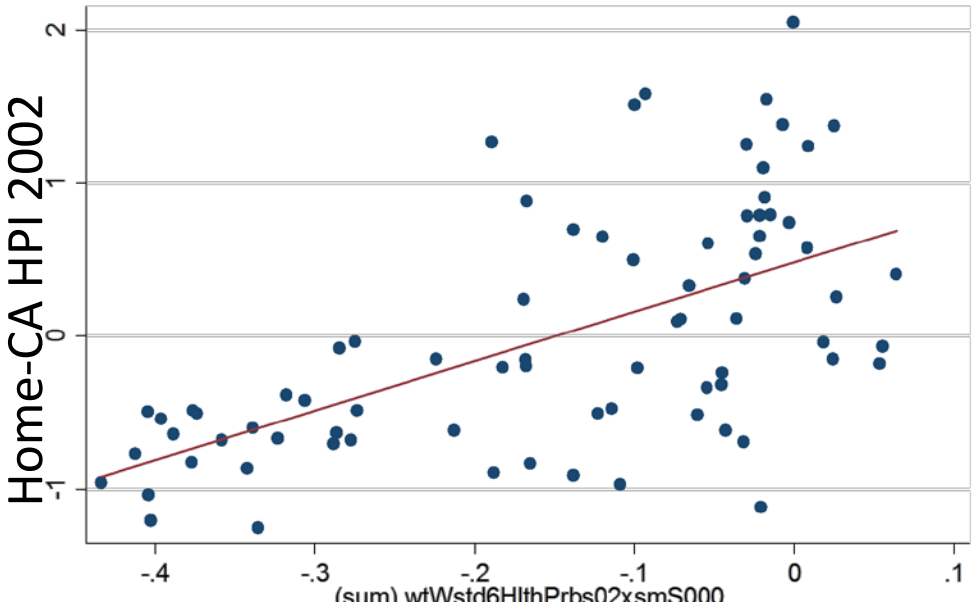
* Symetrized network ties yield the same pattern of results (somewhat stronger coefficients)

Hard challenge : Causal or correlational link?

- How much of the similarity in health risk between two connected communities is due to:
  - Selection (homophily/ exclusion/ discrimination)?
    - e.g., employers in safe communities exclude job applicants from violent areas
  - Causal influence?
    - e.g., safe communities transfer information and resources that increase safety in its connected partners

Implications: Possible Influence through Social Ties





Connected Work-CA HPI 2002
(weighted average per each Home CA)

Homophily in exposure to violence risk
TERGM (Temporal Exponential Random Graph Models):
Ref: Graif et al. 2017 *Social Networks*

Table 1 Dynamic network models (2002–2013): violence effects on tie formation and disolution/pesistence

	Formation	Persistence
+ Network structure		
+ Receiver effects ("Work" community effects)		
Sender effects ("Home" community effects)		
Violent crime rate	-0.75 (2.11)	-7.39 (2.04) ***
Residential stability	0.47 (0.08) ***	-0.05 (0.09)
Racial and ethnic diversity	0.19 (0.06) **	0.14 (0.06) *
Density of local jobs	0.41 (0.2) *	-0.11 (0.17)
Relational effects		
Spatial proximity	1.97 (0.15) ***	0.66 (0.15) ***
Transportation	0.17 (0.03) ***	0.12 (0.03) ***
Dissimilarity		
Violent crime rate	-7.76 (2.21) ***	3.67 (2.47)

Network spillover effects

Table 3. Negative Binomial Regression Predicting 3-Year Spells of Crime by Type, 2004-2006 and 2011-2013

	Overall Crime		Violent Crime		Robbery		Homicide		Property Crime	
	2004-06	2011-13	2004-06	2011-13	2004-06	2011-13	2004-06	2011-13	2004-06	2011-13
Network disadvantage	.162 *	.233 ***	.217 **	.238 ***	.291 **	.278 *	.347 **	.481 ***	.237 ***	.280 ***
	(.068)	(.068)	(.080)	(.070)	(.106)	(.110)	(.124)	(.128)	(.070)	(.074)
Population density	-.365 ***	-.376 ***	-.226 ***	-.218 ***	-.133 †	-.172 †	-.072	-.059	-.436 ***	-.424 ***
	(.050)	(.051)	(.053)	(.054)	(.075)	(.089)	(.095)	(.106)	(.050)	(.049)
Residential stability	-.260 ***	-.272 ***	-.204 ***	-.205 ***	-.203 **	-.172 †	-.030	-.045	-.258 ***	-.230 ***
	(.061)	(.060)	(.057)	(.057)	(.077)	(.094)	(.088)	(.095)	(.065)	(.066)
Ethnic diversity	-.100	-.087	-.123 †	-.092	-.072	-.032	-.069	.025	.004	-.006
	(.069)	(.068)	(.066)	(.065)	(.089)	(.107)	(.101)	(.102)	(.076)	(.072)
Internal Disadvantage	.149 *	.124 †	.303 ***	.277 ***	.147	.156	.355 ***	.388 ***	-.050	-.005
	(.068)	(.067)	(.072)	(.069)	(.093)	(.108)	(.101)	(.100)	(.067)	(.069)
Surrounding disadvantage	.032	.010	.058	.065	.196 †	.184	.099	.070	.081	.037
	(.074)	(.079)	(.077)	(.078)	(.103)	(.129)	(.119)	(.118)	(.074)	(.079)
Surrounding crime	.143 *	.150 **	.152 *	.154 *	.240 **	.139	.114	.166 *	.194 ***	.187 ***
	(.059)	(.057)	(.064)	(.064)	(.089)	(.103)	(.082)	(.074)	(.057)	(.056)
Network crime	.098	.132 *	.016	.059	.028	.269 *	.024	.090	.170 *	.220 **
	(.073)	(.065)	(.078)	(.067)	(.100)	(.117)	(.082)	(.109)	(.074)	(.070)
Temporal lag	-.019	-.005	-.046	-.045	-.025	.005	.037	.017	-.036	-.004
	(.049)	(.050)	(.052)	(.053)	(.069)	(.084)	(.045)	(.048)	(.054)	(.056)
Intercept	8.391 ***	8.066 ***	7.056 ***	6.720 ***	4.925 ***	4.706 ***	1.209 ***	1.137 ***	7.202 ***	7.007 ***
	(.036)	(.037)	(.037)	(.037)	(.049)	(.059)	(.067)	(.070)	(.035)	(.037)
Dispersion Parameter	9.96	9.48	9.62	9.59	5.73	4.05	403.43	403.43	10.58	9.77
	(1.59)	(1.51)	(1.55)	(1.56)	(1.00)	(.70)	(.55)	(.65)	(1.71)	(1.58)
Log Likelihood	-664	-641	-563	-538	-420	-415	-142	-138	-571	-559
AIC	1351	1305	1149	1098	863	852	306	298	1165	1141

NOTES: N =77. Standard Errors in parentheses.

†p<.10; *p < .05; **p < .01; ***p < .001 (two-tailed tests).

Summary Findings: Inter-Neighborhood Ties are Associated with Health Problems

- Communities of similar *socioeconomic and demographic* characteristics are significantly more likely to be connected to each other via commuting flows.

- Communities of similar *health problems* are significantly more likely to be connected to each other via commuting flows.
  - even after controlling for spatial proximity and similarities in socioeconomic characteristics and racial and ethnic composition

- Home-communities with high health problems that are connected to work-communities also with high health problems exhibit a deterioration in health over time
    - Net of geographic proximity and
    - Net of community socioeconomic and demographic characteristics

Related studies (re selection and influence processes):

- Violence levels and similarity in violence shape the likelihood of a connection to exist (be formed or persist) (Graif et al 2017)
- Network disadvantage increases local crime
  - (Graif et al 2019)
  - Violence, homicide, robbery, property crime
  - Strongest association for low income ties
  - Stable connection over time
  - Robust to controlling for internal disadvantage, spatial spillovers of disadvantage and crime, network spillovers of crime

Implications for research and theory

- Re-frames the concept of “neighborhood” as broader spatial and *network context* of routine activities relevant for health behavior,
- Re-frames the idea of spatial regimes under the concept of *inter-neighborhood networks* of influence, structural blocks
- *Neighborhood effects* assessments may need to incorporate residential *and work* neighborhoods: multi-level, cross-level analyses

Policy implications

- The results also inform policy and programs focused on improving the wellbeing of children and communities by highlighting the importance of
  - a) a relational approach to economic and workforce development and
  - b) refining existing thinking on transportation planning, housing policy, and employer - location incentives

Policy Implications and Next Steps

- Contributes to our thinking about opportunities for housing, mobility, and health *interventions* that move beyond a focus on changing the context of residence, focusing also on the neighborhoods of work
  - combating job deserts
  - combating structural isolation (how connected communities are)
  - adjusting connectivity patterns (who are communities connected to)

- Next steps:
  - Dig deeper into network dynamics and the micro level and multi-level processes
  - Thinking deeper about causal connections and mechanisms
  - Multilevel framework of analysis: incorporating Individual level commuting, mobility, and health outcomes.

Contributions and Implications

- Research Contributions
  - Results advance insights into the extra-local spatial and network *mechanisms* that are relevant for health risks
  - Results indicate that mechanisms are not limited within neighborhood boundaries or geographic proximity
- Policy
  - Guide new avenues for *intervention f*ocused on networks, connectivity and *disadvantage* rather than simply local crime
  - A few central neighborhoods in the citywide network of most consequence to others – start there before the remaining neighborhoods
  - *Tax incentives* to have new jobs located in certain neighborhoods first or to encourage employers to hire people from disadvantaged neighborhoods

Next steps: Causal inference solutions?

- Longitudinal data – helps with causal ordering
- Multiple cities – heterogeneity in effects
- Causal modeling,
  - Counterfactual analysis, propensity score matching, inverse probability of treatment weighting
  - (quasi-) Exogenous shocks (external events):
    - Katrina hurricane (flooding effects on individual commuting and neighborhood connectivity)
    - Recession (mass layoffs effects on dissolving ties between two communities)
    - Policy interventions like College Promise, Ban the Box

Supplementary Material

Table 1. Descriptive Statistsics on Nodal Attributes

Variables	Mean	SD	Min	Max
Aggregate Community Area Health Issues				
Birth Problems 2000-2004	0.00	0.94	-1.50	1.62
Infant Mortality Crude Rate	9.23	4.47	1.90	20.60
Pre-Term Births Percent	11.76	2.90	7.30	17.60
Low Birth Weight Percent	10.23	3.45	5.50	17.40
Very Low Birth Weight Percent	2.12	0.95	0.60	4.00
Birth Problems 2005-2009	0.00	0.96	-1.17	1.92
Infant Mortality Crude Rate	8.57	4.41	1.50	22.60
Pre-Term Births Percent	11.86	2.75	8.00	16.40
Low Birth Weight Percent	10.40	3.49	6.30	16.70
Very Low Birth Weight Percent	2.14	1.02	0.80	4.60
Birth Problems 2010-2014	0.00	0.94	-1.20	2.47
Infant Mortality Crude Rate	8.12	4.42	1.40	24.50
Pre-Term Births Percent	10.99	2.47	7.40	17.80
Low Birth Weight Percent	10.16	3.55	6.00	23.90
Very Low Birth Weight Percent	2.06	0.98	0.50	5.00
2000 Community Area Census Data				
Diversity	0.12	1.04	-1.54	2.40
Disadvantage	-0.04	0.91	-1.24	2.38
Stability	-0.01	0.97	-2.11	1.73

Table 2. Network Structure Desctiptive Statistics												
Characteristics	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Number of Nodes	77	77	77	77	77	77	77	77	77	77	77	77
Number of Edges	639	583	582	624	559	529	606	530	540	551	504	525
Min Indegree	0	0	0	0	0	0	0	0	0	0	0	0
Median Indegree	2	2	1	2	1	1	1	1	0	1	1	0
Max Indegree	76	76	76	76	76	76	76	76	76	76	76	76
Min Outdegree	2	2	2	2	1	3	4	3	3	3	3	3
Median Outdegree	8	7	7	8	7	6	7	6	7	7	7	7
Max Outdegree	14	14	13	14	13	15	13	18	14	13	14	15
Density	0.109	0.0996	0.099	0.107	0.096	0.09	0.104	0.091	0.092	0.094	0.086	0.0897
Transitivity	0.677	0.683	0.665	0.632	0.5998	0.616	0.6697	0.595	0.601	0.645	0.628	0.6497
Number of Mutual Ties	63	56	52	61	55	39	48	43	46	41	46	45
Dyadic Ratio	0.825	0.839	0.837	0.828	0.847	0.846	0.826	0.848	0.847	0.8397	0.859	0.851
Edgwise Ratio	0.197	0.192	0.179	0.196	0.197	0.147	0.158	0.162	0.17	0.1488	0.183	0.171
Note: These descriptive statistics are from the network using a 5% cutoff												

References

- Graif, Corina, Brittany N. Freelin*, Yu-Hsuan Kuo*, Hongjian Wang*, Zhenhui Li, and Daniel Kifer. 2019. "Network spillovers and neighborhood crime: A computational statistics analysis of employment based networks of neighborhoods Forthcoming. *Justice Quarterly.* (online ahead of print)
- Wang, Hongjian*, Huaxiu Yao*, Daniel Kifer, Corina Graif, and Zhenhui Li. *2018. Non-Stationary Model for Crime Rate Inference Using Modern Urban Data. IEEE Transactions on Big Data.*
- Graif, Corina, Alina Lungenu, and Alyssa Yetter*. 2017. "Neighborhood isolation in a Rust Belt city: Violent crime effects on increasing structural isolation and homophily in inter- neighborhood commuting networks" *Social Networks. 51:40-59.*
- Arcaya, Mariana, Corina Graif, Mary Waters, and S.V. Subramanian. 2016. "Health Selection into Neighborhoods among Moving to Opportunity Families." *American Journal of Epidemiology*. 183(2):130-7.
- Wang, Hongjian*, Daniel Kifer, Corina Graif, and Zhenhui Li. 2016. "Crime Rate Inference with Big Data." 2016. *Proceedings of the Knowledge Data Discovery ACM SIGKDD*.
- Graif, Corina and Stephen Matthews. 2017. "The Long Arm of Poverty: Extended Geography of Neighborhood Effects on Child Victimization" *Justice Quarterly. 34(6):1096-1125.*
- Graif, Corina. 2011. *Mobility in Isolation*. Doctoral Thesis. Harvard University

Dynamic Social Network Models

TERGM models – different crime types.

Formation models	Model 1a – Homicides		Model 2a – Robbery		Model 3a – Robbery and aggravated assault	
Network structure						
Edge	−4.07 (0.17)	***	−4.01 (0.17)	***	−3.84 (0.19)	***
Reciprocity	0.05 (0.17)		0 (0.18)		−0.02 (0.19)	
Geometrical weighted in-degree (popularity spread)	−4.07 (0.28)	***	−4.09 (0.26)	***	−4.06 (0.27)	***
Receiver effects ("Work" community effects)						
Violent Crime Rate	0.48 (0.33)		0.03 (0.01)	†	0.15 (0.1)	
Residential stability	−0.84 (0.07)	***	−0.83 (0.07)	***	−0.82 (0.08)	***
Racial and ethnic diversity	−0.29 (0.05)	***	−0.29 (0.05)	***	−0.32 (0.06)	***
Density of local jobs	1.29 (0.21)	***	1.28 (0.21)	***	1.15 (0.23)	***
Sender effects ("Home" community effects)						
Violent Crime Rate	−0.9 (0.36)	*	−0.04 (0.02)	*	−0.27 (0.12)	*
Residential stability	0.47 (0.08)	***	0.45 (0.08)	***	0.45 (0.08)	***
Racial and ethnic diversity	0.18 (0.06)	**	0.14 (0.06)	*	0.14 (0.06)	*
Density of local jobs	0.31 (0.21)		0.32 (0.21)		0.25 (0.24)	
Relational effects						
Spatial proximity	1.98 (0.15)	***	1.98 (0.15)	***	2.01 (0.15)	***
Transportation	0.19 (0.03)	***	0.18 (0.03)	***	0.18 (0.03)	***
Dissimilarity						
Violent Crime Rate	−0.37 (0.39)		−0.05 (0.02)	*	−0.4 (0.13)	**
Residential stability	−0.29 (0.09)	***	−0.29 (0.08)	***	−0.28 (0.09)	**
Racial and ethnic diversity	−0.21 (0.06)	***	−0.18 (0.06)	**	−0.21 (0.06)	***
Density of local jobs	−0.25 (0.21)		−0.21 (0.21)		−0.14 (0.24)	
AIC	−741,410		−741,413		−593,211	
BIC	−741,259		−741,262		−593,062	

Overall Crime

Table C1. Negative Binomial Regression Predicting 3-Year Aggregate All Crime, Low Income Ties (N=77)

	2004-2006				2007-2009	2010-2012	2011-2013
	M1	**M2**	**M3**	**M4**	**M4**	**M4**	**M4**
Network disadvantage	.455 ***	.248 ***	.208 **	.161 *	.228 ***	.241 ***	.233 ***
	(.059)	(.064)	(.074)	(.068)	(.066)	(.062)	(.068)
Population density		-.311 ***	-.310 ***	-.372 ***	-.370 ***	-.359 ***	-.377 ***
		(.051)	(.050)	(.047)	(.045)	(.044)	(.048)
Residential stability		-.361 ***	-.348 ***	-.255 ***	-.281 ***	-.244 ***	-.270 ***
		(.054)	(.055)	(.060)	(.058)	(.056)	(.059)
Ethnic diversity		-.212 **	-.190 **	-.092	-.105 †	-.102 †	-.086
		(.066)	(.069)	(.066)	(.062)	(.058)	(.065)
Disadvantage index		.145 *	.115	.145 *	.125 †	.145 *	.123 †
		(.069)	(.074)	(.067)	(.065)	(.060)	(.067)
Spatial lag disadvantage			.089	.036	-.028	-.039	.011
			(.086)	(.074)	(.074)	(.074)	(.079)
Spatial lag crime				.134 *	.173 ***	.139 **	.148 **
				(.054)	(.046)	(.044)	(.052)
Network crime				.103	.072	.156 **	.133 *
				(.071)	(.057)	(.053)	(.064)
Intercept	8.481 ***	8.405 ***	8.404 ***	8.391 ***	8.294 ***	8.119 ***	8.066 ***
	(.059)	(.041)	(.040)	(.036)	(.035)	(.034)	(.037)
Dispersion Parameter	3.703	7.872	7.979	9.945	10.672	11.375	9.477
	(.573)	(1.247)	(1.264)	(1.584)	(1.702)	(1.818)	(1.512)
Log Likelihood	-705	-674	-673	-664	-654	-639	-641
AIC	1415	1362	1363	1349	1329	1298	1303

Standard Errors in parentheses

†p<.1 *p < .05 **p < .01 ***p < .001 (two-tailed)

Detroit and Arlington (Study under review)

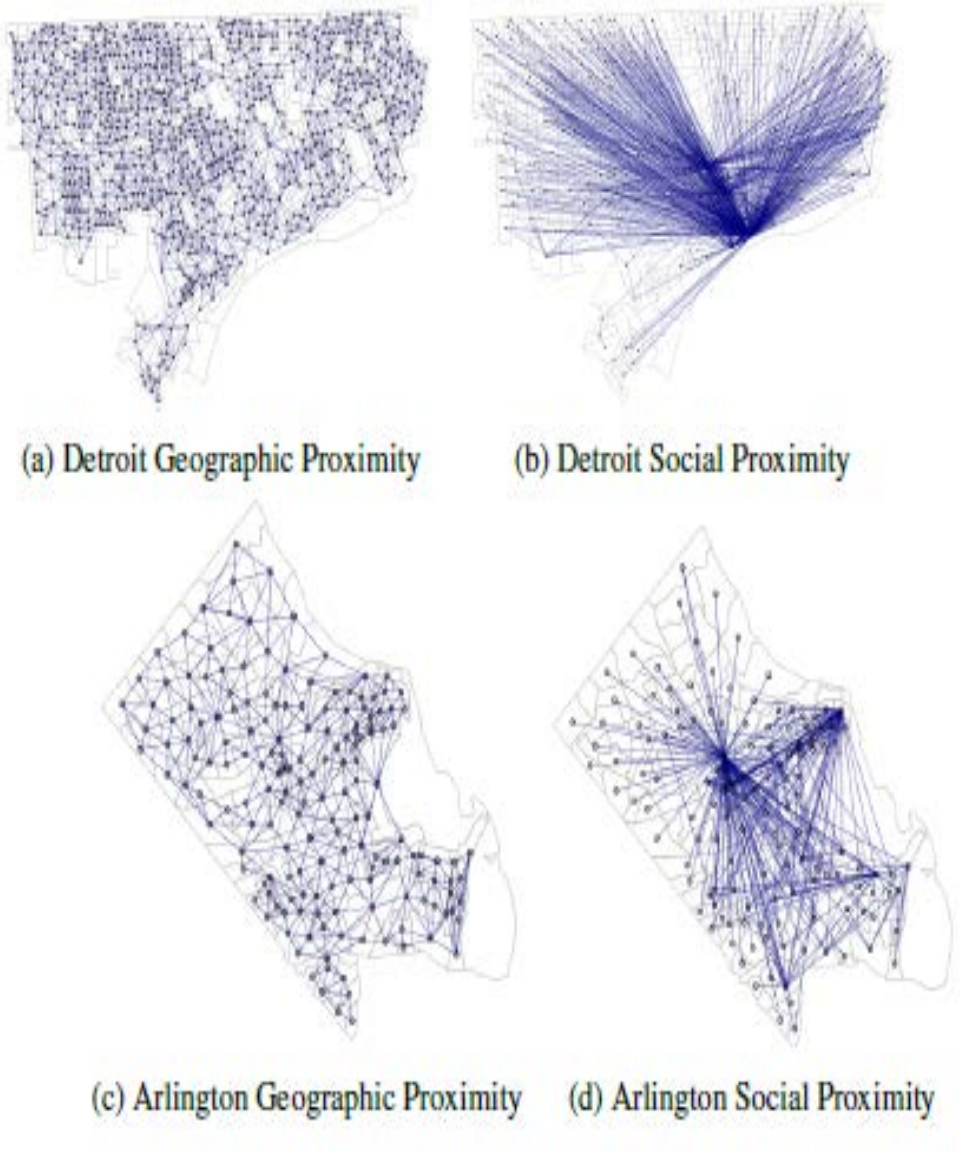
Detroit

Variable	Geographic	Social	Geo. & Soc.
Number of regions	871	826	871
Number of nonzero links	4,972	7,322	12,006
Percentage nonzero weights	0.65	1.07	1.58
Average number of weights per unit	5.71	8.86	13.78

Arlington

Variable	Geographic	Social	Geo. & Soc.
Number of regions	173	172	173
Number of nonzero links	982	1,758	2,546
Percentage nonzero weights	3.28	5.94	8.51
Average number of weights per unit	5.68	10.22	14.72

Table 1: Descriptive Statistics of Neighborhood Structures



(a) Detroit Geographic Proximity (b) Detroit Social Proximity

(c) Arlington Geographic Proximity (d) Arlington Social Proximity

Figure 1: Plot of Neighborhood Structure