

Capturing Context within a Region and its Communities: Applications of Typologies and Cluster Analyses

Joanna Ganning

Hua Qin

Courtney Flint

Stephen Gasteyer

University of Illinois Urbana-Champaign

August 4, 2007

Background

- Typology research is used to
 - **Contextualize regions**
 - Measure intervention effects
 - Explore causal mechanisms, hypothesis testing
 - Explore group potential
 - Tailor policies to specific contexts

Background

- Traditional quantitative methods tend to obscure heterogeneity rather than highlight it.
- The vast majority of quantitatively-oriented community studies use traditional quantitative methods.
- Between 2001 and 2003 just 1-5% of articles published in *American Journal of Community Psychology* used methods that highlighted heterogeneity. (Luke 2005)

Background

- Methods to contextualize study regions include
 - cluster analysis
 - social network analysis
 - hierarchical linear modeling
 - geographic information systems

Approach

- This paper illustrates that the use of *multiple approaches* to creating clusters and typologies allows for the contextualized study of regions *at multiple scales*
- Part of a larger project assessing the Southern Illinois region, particularly the role of a university-funded agricultural research center.

Theoretical Background

- Typologies are not new.
 - Classic sociology typologies include Tonnie's *Gemeinschaft and Gesellschaft* and Durkheim's Mechanical and Organic Solidarity.
- Methods are not new.
 - McKinney (1966) outlines an 8-step implementation guide
 - Aldenderfer and Blashfield (1984) provide a guide to reporting cluster analysis findings

Southern Illinois Typologies

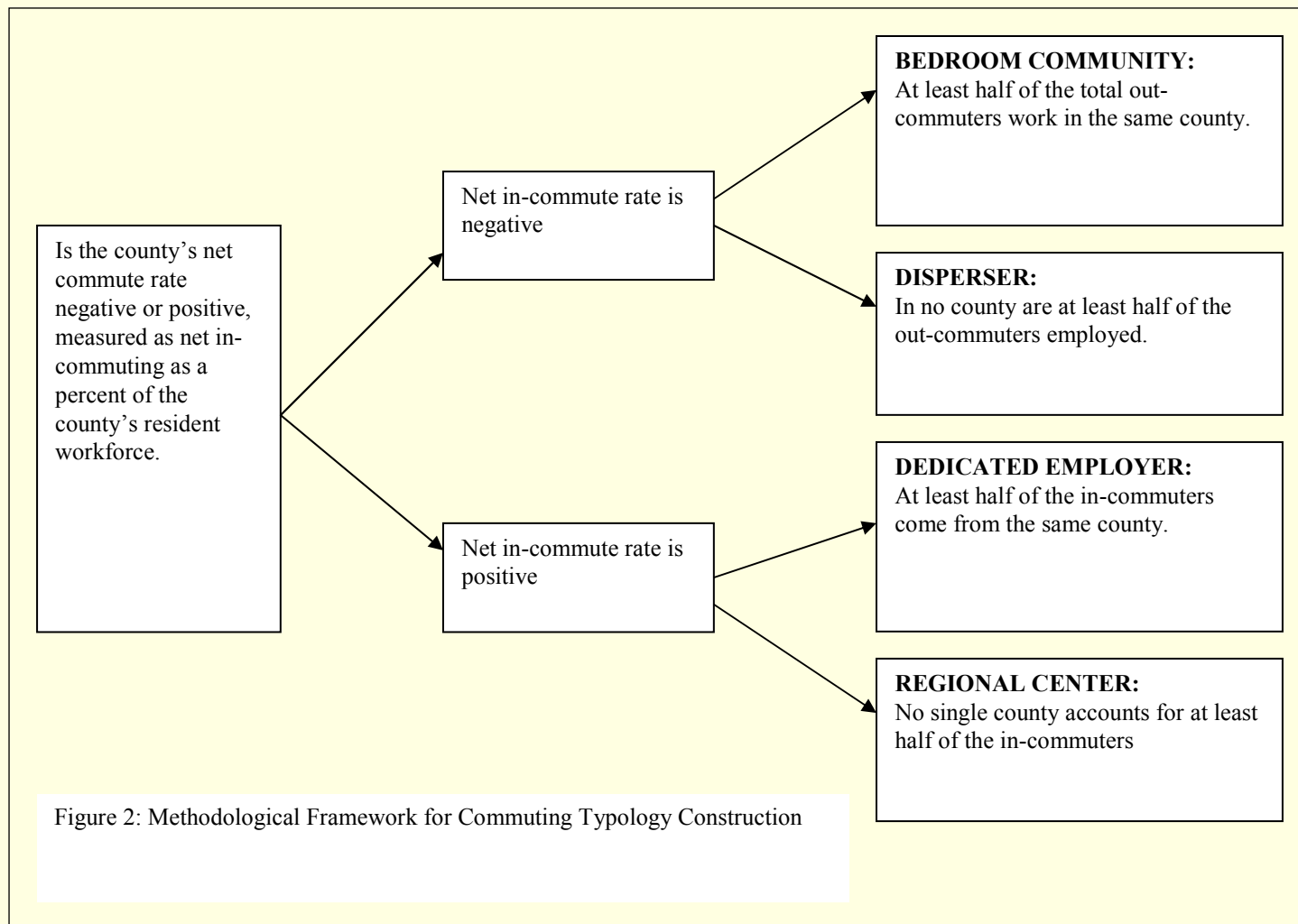
1. Simple Typologies

- a) Commuting
- b) Land Use

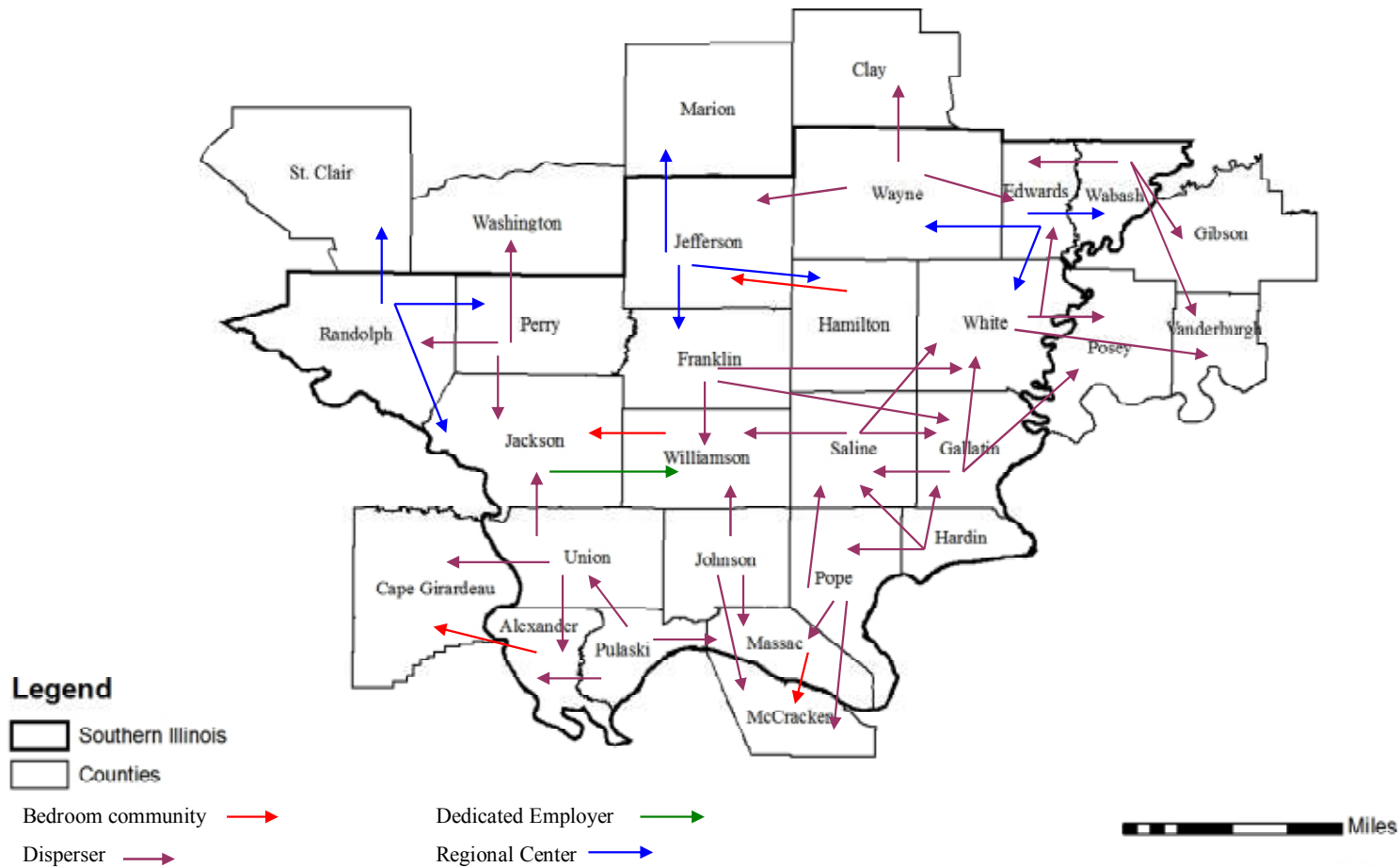
1. Multi-dimensional Typologies

- a) County-level using hierarchical clustering
- b) Census tract-level using Factor Analysis and c-means fuzzy-set clustering

Typology 1: Commuting



Commuting Results



Data sources: US Census Bureau, Journey to Work, 2000

Ganning, Qin, Flint, Gasteyer 2007

Typology 2: County-level Land Use



- Land and resource intensive industries, specifically ag, forestry, and mining, have a strong presence in Southern Illinois.
- Focusing on land use variables is thus a key part of understanding community dynamics in Southern Illinois.

County-level Land Use

		Mining employment, measured by location quotient relative to US		
		Non-Mining Counties (location quotient < 2)	Mining Counties (location quotient > 2)	
			coal	non-coal
Farm Employment, measured by location quotient relative to US	Average (location quotient 0-2)	Jackson, Williamson		
	Above average (location quotient 2-5)	Edwards, Massac, Jefferson	Wabash, White, Franklin, Perry, Randolph, Saline	Alexander
	High (location quotient >5)	Pope, Hamilton	Gallatin	Hardin, Johnson, Union, Pulaski, Wayne

Typology 3: County-Level Multi-Dimensional

- Standard and practical unit of analysis for regional studies
- Often the county is the lowest geographical level at which data is available
- Key scale of governance and identity



County-Level Multi-Dimensional Approach

Counties analyzed across 9 variables:

- Population density
- In-migration rate
- Per capital personal income
- Poverty rate
- Percent of the population (ages 25 and over) having at least a high school diploma
- Crime rate
- Percent of land cover in agricultural use
- Total air pollutant emissions density
- Percent of surface water area threatened

County-Level Clustering Results

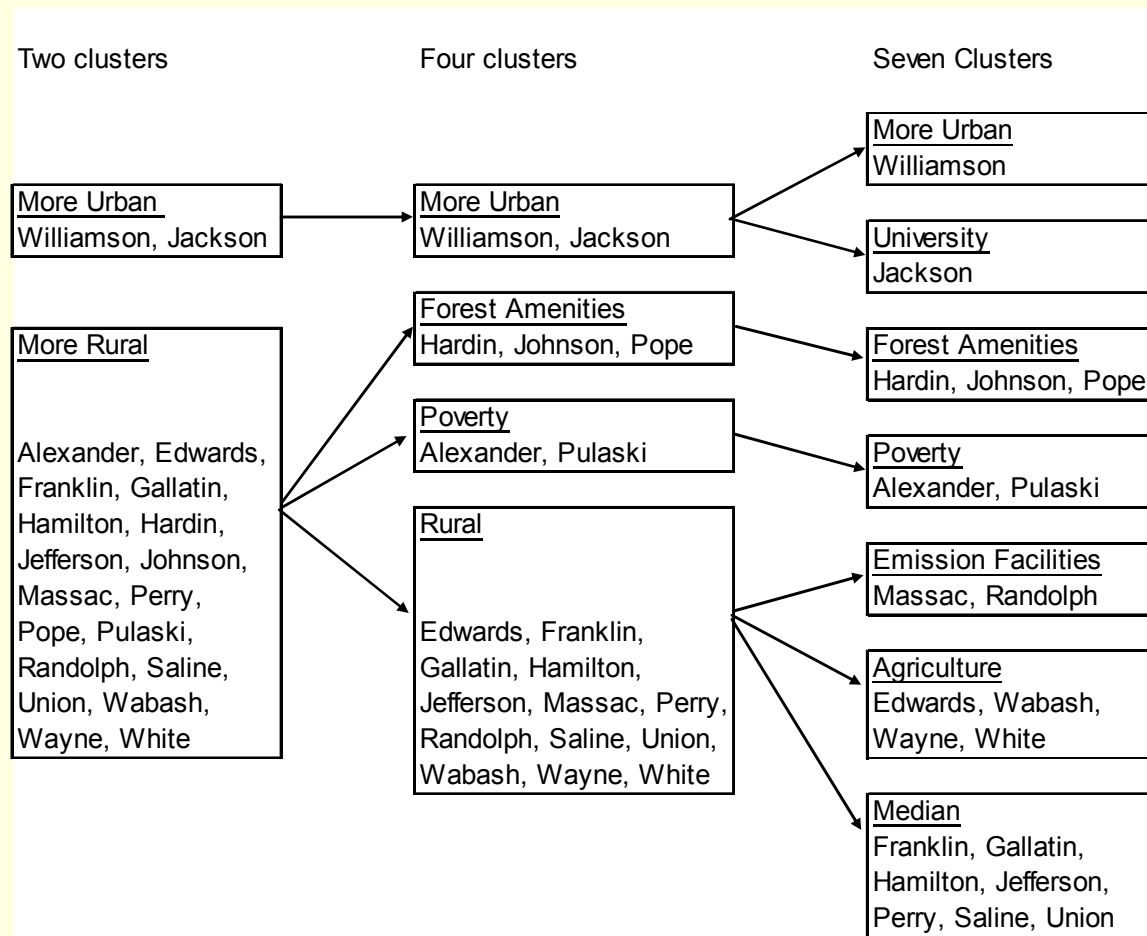


Figure 4: County-level cluster map of southern Illinois

Typology 4: Multidimensional Census tract-Level



- Two primary motivations:
 - Greater statistical power with larger number of observations
 - A number of population processes occur at a sub-county level

Tract-Level Factor Analysis Results

	Component			
	Fitchen-Style Migration	Education & Housing Costs/Age	Ag Employment & Mobile Homes	Seasonal Housing & Plumbing
POPULATION DENSITY	0.632	-0.042	-0.238	-0.204
PER CAPITA PERSONAL INCOME	-0.692	0.546	-0.231	0.001
POVERTY	0.944	-0.070	0.024	0.036
HOMEOWNERSHIP	-0.913	-0.122	0.191	0.135
PERCENT IN-MIGRANTS	0.879	0.183	-0.231	-0.028
GED PLUS	0.150	0.703	0.004	-0.238
MEDIAN HOME AGE	0.022	-0.669	-0.544	-0.237
MEDIAN RENT	0.158	0.682	-0.453	-0.143
MEDIAN OWNER-OCCUPIED HOUSING VALUE	-0.137	0.914	0.049	0.005
PERCENT MOBILE HOMES	-0.043	0.067	0.822	0.224
AG. EMPLOYMENT	-0.347	-0.120	0.707	0.146
SEASONAL VACANCY	-0.208	0.018	0.225	0.806
INCOMPLETE PLUMBING	-0.076	-0.132	0.285	0.749
UNEMPLOYMENT	0.382	-0.406	-0.276	0.494
<i>eigenvalue</i>	4.344	3.009	2.014	1.000

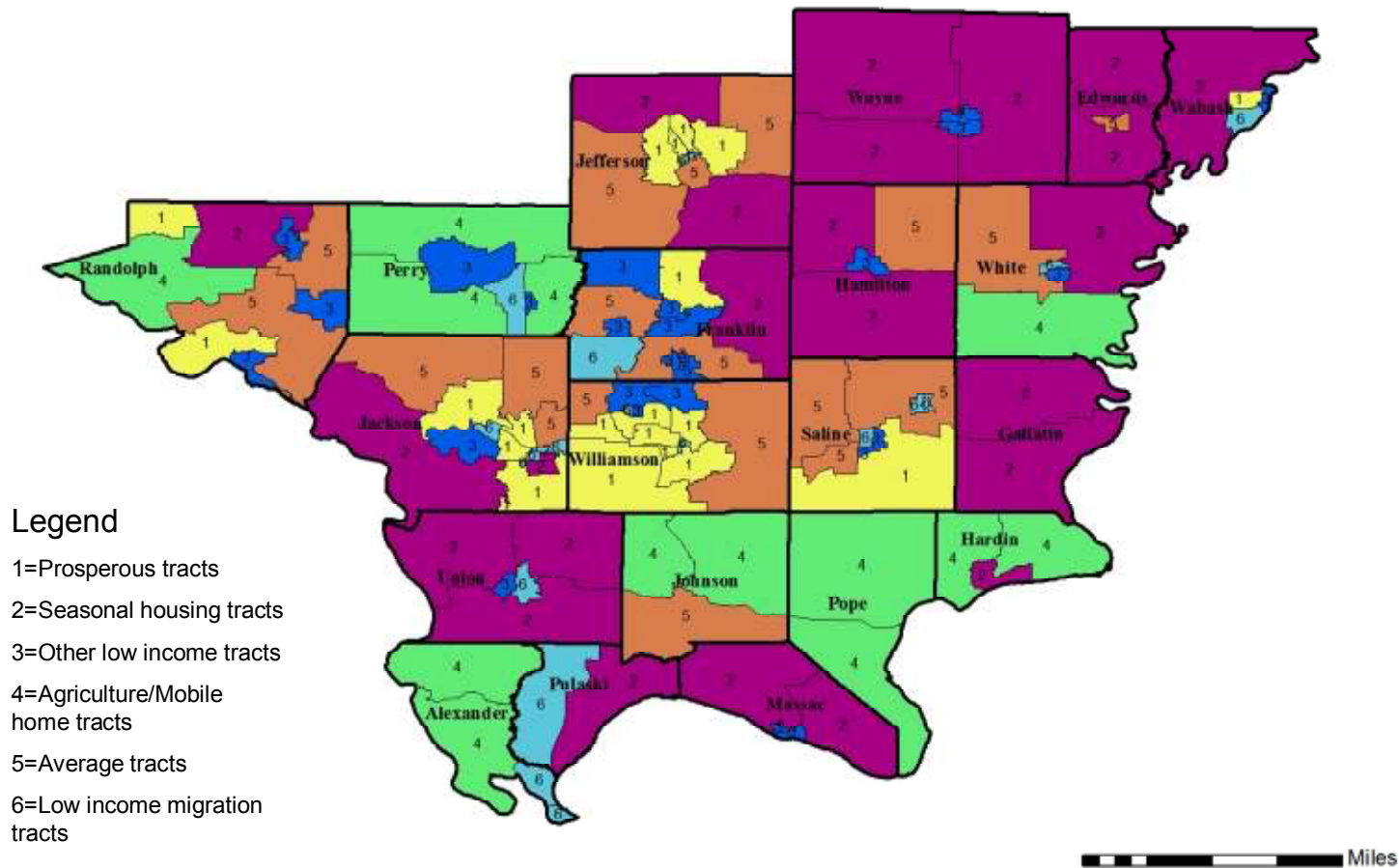
Data source: US Census Bureau, Census 2000. Except where otherwise noted variables follow the Census Bureau's definitions. "GED Plus" refers to the percent of the population ages 25 and over who have at least a high school diploma or equivalency. Seasonal vacancy refers to the percent of all housing units that are vacant for seasonal use. "Percent mobile homes" indicates the percent of all housing units which are mobile homes. "Incomplete plumbing" measures the percent of all housing units which lack complete plumbing facilities. Percent in-migrants is measured as the percent of persons who lived outside the county in 1995 and in 2000 lived in the county, divided by the 2000 county population.

Tract-level Fuzzy-set Cluster Results

A c-means fuzzy-set cluster analysis on the Factor Analysis scores produced 6 meaningful clusters:

- Prosperous Clusters
- Seasonal Housing Clusters
- Other Low Income Clusters
- Ag/Mobile Home Clusters
- Average Clusters
- Low Income Migration Clusters

Tract-Level Clustering Results



Synthesis

Local areas do not behave similarly from typology to typology.

- The southeastern part of the region was:
 - Uniformly “Disperser” in the commuting typology
 - “Emissions,” “Forest Amenities,” or “Median” counties in the county-level multi-dimensional
 - “Ag/mobile home” cluster, “Median” cluster, or “Seasonal housing” cluster in the Census tract-level typology.

Synthesis

This is not inconsistent with expectations

- Counties that rely on forest amenities are unlikely to have a large employment base at home.
- An overlap between “forest amenities” (county-level typology) and “seasonal housing/incomplete plumbing” (tract-level typology) is not surprising. Areas with outdoor amenities tend to attract seasonal residents.



Synthesis

- Randolph, Jefferson, and Edwards Counties are the only Regional Centers in the Commuting Typology, yet seem dissimilar in the other typologies.
- Paths taken to a strong employment base vary based on history of county. This is consistent with theory and expectations.



Synthesis

Not all agriculture is the same

- In county-level multi-dimensional, northeastern counties are most agricultural and have high incomes.
- In Census tract-level, southeastern tracts are most agricultural and have low incomes.
- Use two different measures, land cover and employment. NE has large farms with low employment, while SE has smaller farms with larger employment relative to county.

Conclusion

- Using existing data, approaches, and methods, typologies can create a level of contextualization infrequently seen in community studies.
- By creating multiple levels of comparison, stemming from different perspectives, local areas' features can be integrated into an understanding of the region.