

# Characterizing the Regional Structure in the United States: A county-based analysis of labor market centrality

Nikhil Kaza Ph.D.  
Department of City & Regional Planning,  
Environment, Ecology & Energy Program.  
University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3140  
[nkaza@unc.edu](mailto:nkaza@unc.edu)

Katherine Nesse Ph.D.  
City of Bellevue, WA 98004  
[knesse@bellevuewa.org](mailto:knesse@bellevuewa.org)

---

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11

12

# 13 Characterizing the Regional Structure in the United States: A 14 county-based analysis of labor market centrality

---

15

## **Abstract**

16

17

18

19

20

21

22

23

24

25

Categorizing places based on their network connections to other places in the region reveals not only population concentration but also economic dynamics that are missed in other typologies. The US Office of Management and Budget categorization of counties into metropolitan/micropolitan and central/outlying is widely seen as insufficient for many analytic purposes. In this paper we use a coreness index from network analysis to identify labor market centrality of a county. We use county-to-county commute flows, including internal commuting, to identify regional hierarchies. Indicators broken down by this typology reveal counterintuitive results in many cases. Not all strong core counties have large populations or high levels of urbanisation. Employment in these strong core counties grew faster in the post-recession (2008-2015) than in other types of counties. This economic dimension is missed by other typologies, suggesting that our categorization may be useful for regional analysis and policy.

## 26 1. Introduction

27

28 Researchers and policy-makers recognized that major cities are labor market centers that draw from their  
 29 peripheral regions. For a while, states, local and federal agencies, counties and other entities created their  
 30 own ad hoc definitions of labor market areas based on these centers (Congressional Research Service  
 31 2014). In 1949, the US Office of Management and Budget created a standard definition of these labor  
 32 market areas based on commuting patterns and called them “Standard Metropolitan Areas” (Klove  
 33 1952). The standardization of these areas enabled clearer communication between governments and  
 34 public agencies. These definitions have been updated approximately once a decade since then. In 2015,  
 35 the Office of Management and Budget (OMB) delineated 945 Core Based Statistical Areas (CBSA) in  
 36 the United States and Puerto Rico. Each of these CBSAs are collections of counties. Of these, 389 are  
 37 Metropolitan (MSA), with core areas of 50,000 or more people, and the rest are Micropolitan ( $\mu$ SA),  
 38 with core areas of 10,000 to 49,999 people\*. For very large MSAs, such as New York, OMB has created  
 39 Metropolitan Divisions that delineate smaller labor markets within the MSA (Congressional Research  
 40 Service 2014).

41

42 OMB classifies two types of counties within the statistical areas: Central, those counties that contain all  
 43 or a substantial portion of the urbanized area; and Outlying, those counties that have employment  
 44 interchange measure with the Central counties above 25%. In other words, the centrality of the county is  
 45 defined by the urban population attributes of the county rather than its relative location in the commuting  
 46 network. In 2015, a vast majority of the counties within CBSA are considered central; only 29% of the  
 47 counties are peripheral/outlying (see Table 1). This is even more stark within  $\mu$ SAs where only 14% are  
 48 considered peripheral. CBSAs are dominated by the central counties. They account for 92.5% of the  
 49 CBSA population. These central counties are crucial to the delineation of these statistical regions and  
 50 encompass the economic core of the country.

51 Table 1 Types of CBSAs and Counties in Conterminous United States. Source: OMB (2015)

CBSA Type	County Type		Total
	Central	Outlying	
MSA	725	436	1,161
$\mu$ SA	560	94	654
Total	1,285	530	1,815

\* In the rest of the paper, for the sake of simplicity we refer only to conterminous United States excluding the states of Hawaii, Alaska and territories such as Puerto Rico and Guam.

52

53 While statistical areas are the most commonly used way of delineating labor market areas, several  
54 researchers have found the central/outlying/non-metropolitan categorization to be too crude to describe  
55 the diversity of counties in the United States (e.g. Isserman 2005; Waldorf and Kim 2018; Wang et al.  
56 2012). Many researchers have proposed their own typologies, though still based on commuting patterns  
57 or population levels, with the objective to better understand labor market areas (e.g. Fowler, Jensen, and  
58 Rhubart 2018; Han and Goetz 2019; Tolbert and Sizer 1996) or to better align public programs (e.g.  
59 Cook and Mizer 1994; De Lew 1992; Hewitt 1989; Lipscomb and Kashbrasiev 2008).

60

61 There is often a conflation of urbanicity with metropolitan areas. Isserman (2005) identified the  
62 differences between the US Census demarcation of urban and rural and OMB delineation of metropolitan  
63 and non-metropolitan areas, even when they are frequently used interchangeably. The former is about  
64 *separation* of densely built from sparsely built places, while the latter is primarily about *integration* of  
65 residence and place-of-work. Consequently, OMB's metropolitan areas include large swaths of rural  
66 lands, centered on urban counties. Nonetheless, the context of the metropolitan area becomes important  
67 feature for determination of the urbanity of a place. For example, the Economic Research Service (ERS)  
68 of the US Department of Agriculture created Rural-Urban Continuum Codes (RUCC) to distinguish  
69 counties based on the population size of the metro area and proximity to metro areas for a total of nine  
70 categories (3 metro and 6 non-metro)(ERS 2013). ERS has created a commuting rubric at the sub-county  
71 level as well. The Rural-Urban Community Area (RUCA) codes identify census tracts as metropolitan,  
72 micropolitan or nonmetropolitan and breaks these down based on the size of the commute flow (ERS  
73 2019). This 10-category classification system is more detailed both geographically and categorically but  
74 is still based in the OMB system. The National Center for Health Statistics (NCHS) created their own  
75 urban-rural typology of counties (4 metro and 2 non-metro) skewed towards metropolitan areas, by  
76 arguing that they stand in for urban and rural distinctions (NCHS 2014). These categories also include  
77 the distinctions between central and fringe counties in large metro regions as well as based on the size of  
78 the region (large, medium, small). For example, Han and Goetz (2015) argue that resilience patterns are  
79 different across different types of counties. Using the economic recession of 2008 and the recovery  
80 pattern as evidenced by employment changes, they argue that counties with RUCC codes 1-5 (large  
81 population metro and non-metro areas) have more resistance rather than resilience. Interestingly, an  
82 USDA report conflates rural with non-metropolitan areas and argues that rural areas are slow to recover  
83 post-recession (Farrigan 2019). These concepts of centrality, urbanity and proximity, condition our  
84 understanding of disparities and challenges faced by different regions (e.g. Cutter, Ash, and Emrich

85 2016; Ingram and Franco 2012; Scala and Johnson 2017). While these are important, we argue that they  
86 are complementary and not interchangeable with centrality.

87

88 In this paper, we seek to establish a different notion of centrality based on the position of the county in a  
89 commuting network. We aim to uncover the regional structure by explicitly focusing on the network  
90 rather than node attributes (unlike OMB's characterization of centrality). In particular, counties (or  
91 equivalents) that have small population (either because of constricted boundaries or because of a lack of  
92 residential lands) might still be destination areas for commuting. On the other hand, counties that have  
93 large populations may not necessarily have sufficient economic activity to justify the Central  
94 designation. By focusing on the relationships among counties as a network, rather than the county  
95 attributes, we can begin to uncover some regional structures such as integration and core-periphery  
96 structures. The regions may be (multi) core based, or comprise exclusively of peripheral nodes. Our aim  
97 is to demonstrate that such understanding complements our conventional understanding of centrality  
98 used in the literature and policy analysis. We also seek to demonstrate meso-level regional structure  
99 rather than intra-regional spatial structure using a national analysis. We demonstrate that the  
100 positionality-based (as opposed to attribute based) centrality is correlated with economic growth and  
101 sectoral specialization. We also demonstrate that centrality is not correlated, in particular, with urbanity  
102 or population levels.

103

## 104 **2. Network Analyses**

105

106 Network analyses in regional science have a long history, though there have been significant divergences  
107 between geographers and network scientists (for an extensive literature review, see Ducruet and  
108 Beauguitte 2014). For example, Nystuen and Dacey (1961) use graph theory and commuting flows to  
109 identify regional hierarchies and nodal regions. Tong and Plane (2014) use spatial optimization  
110 techniques on the commuting network of all inter-county commuting linkages to identify clusters that  
111 rival OMB's Metropolitan area delineations. Nelson and Rae (2016) use community detection  
112 techniques to derive the mega-region structures. Using different community detection techniques and  
113 statistical inference, He et al. (2020) propose that there are multiple overlapping regions in the United  
114 States, hitherto unrecognized by the OMB or the other delineations such as megaregions. These are but  
115 a few examples of use of graph theoretical analyses applied to regional science problems. Many of them  
116 rely on clustering of sub-geographies to construct a larger geographical region. Among network science

117 approaches in regional science, to our knowledge, very few have focused on centrality (see e.g. Neal  
118 2011; Sigler and Martinus 2016).

119

120 Centrality is a property of a node in a network indicating its relative importance (Degenne and Frose  
121 1999). Many such measures of centrality have been proposed including degree centrality, betweenness  
122 centrality and eigen centrality (e.g. Bonacich 1972; Freeman 1978). For a network with nodes and edges  
123 representing the connection between the nodes, a node degree is equal to the number of edges that are  
124 incident on the node (or conversely number of nodes it is connected to). A node with a higher degree is  
125 considered more central than ones with lower degree. However, degree centrality is a local measure,  
126 ignoring its importance in the overall network, through indirect connections. Betweenness centrality, a  
127 global measure, is based on how frequently a node appears on a path between two other nodes. While  
128 degree centrality is a measure of number of walks of length one that the node appears in, eigen centrality  
129 generalizes it to the number of walks of infinite length (Newman 2016). These centrality measures have  
130 been used to study many geographic networks such as transit systems (Derrible 2012), street networks  
131 (Agryzkov et al. 2019; Kirkley et al. 2018), knowledge networks (Maggioni and Uberti 2009) and global  
132 value chains (Cingolani, Panzarasa, and Tajoli 2017).

133

134 While these centrality measures are important characterizations of nodes, they are not always appropriate  
135 to understand the regional structure. In a commuting network among counties, betweenness centrality  
136 does not make intuitive sense. The degree centrality is a measure of how many commuters are incident  
137 on a county, while eigen centrality is about how well the county is connected to other central counties.  
138 While these two make some sense for commuting networks, they do not shed light on meso-level  
139 structural properties of the network. There is another concept called  $k$ -core centrality that relies on  
140 successive pruning of a network to identify more closely connected nodes (Seidman 1983). These  
141 concepts of identifying core and periphery of a network has found many applications ranging from  
142 airport commuting networks (Verma, Araújo, and Herrmann 2014) to gene regulatory networks (Narang  
143 et al. 2015). In this paper, we focus on this measure.

144

145

146

147

### 148 3. Methods & Data

#### 149 3.1. *s*-core decomposition of a graph

150

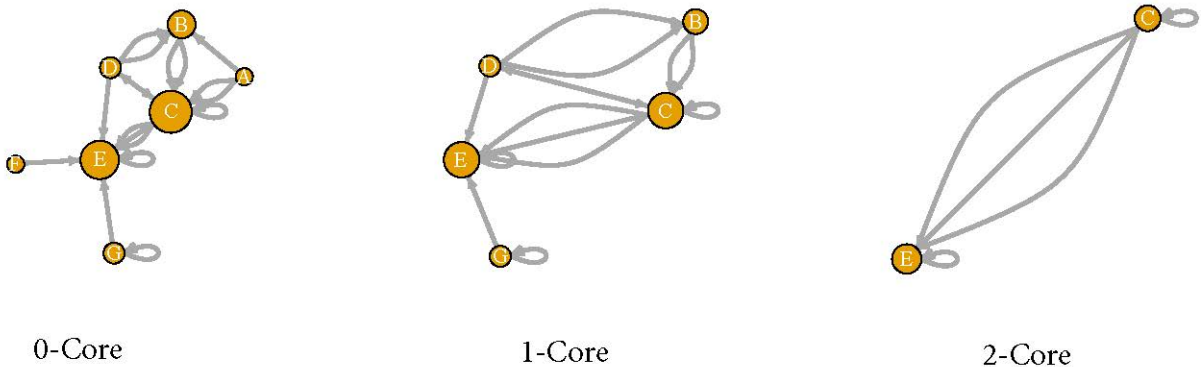
151 A  $k$ -core of an unweighted simple binary graph is its subgraph where all the nodes have at least degree  $k$   
 152 (Dorogovtsev, Goltsev, and Mendes 2006; Seidman 1983). This subgraph is obtained by iteratively  
 153 removing nodes from the network whose degree is less than  $k$  until a stable set of vertices with the  
 154 minimum degree is reached. A node in a network has a coreness index  $k$ , if it belongs to a  $k$ -core but not  
 155 a  $k + 1$ -core. Algorithms to calculate these indices quickly are proposed by Batagelj and Zaveršnik  
 156 (2011).

157 This can be generalized to a directed network by focusing on the indegree; i.e. a  $k$ -core is the subgraph,  
 158 where all nodes have an in-degree  $k$ . We can also generalize this concept to a weighted graph by using  
 159  $s$ -core decomposition, where degree of the vertex is replaced by strength of the vertices (Eidsaa and  
 160 Almaas 2013). If, the edge weight between nodes  $i$  and  $j$  is denoted by a non-negative  $w_{ij}$ , then the  
 161 strength of the vertex  $i$  is defined as

$$162 \quad s_i \equiv \sum_{j \in N_i^-} w_{ij}$$

163 where  $N_i^-$  is the in-neighborhood of  $i$ . The  $s$ -core is a subgraph where the nodes have at least strength  
 164  $s$ . As long as  $w_{ij} \in \mathbb{Z}^+$ , we can replace an edge in the graph with  $w_{ij}$  multi-edges, and the decomposition  
 165 of the graph by strength and degree are equivalent.

166 Figure 1 Illustration of network decomposition into core and periphery. Vertices are sized based on their in-degree.



167

168

169 The  $s$ -core decomposition is illustrated in Figure 1 for a directed graph with multiple edges including  
 170 loops. The entire graph in the figure is part of 0-core. Nodes A and F have in-degree 0, and therefore are  
 171 not part of the 1-core of the graph (subgraph induced by nodes B, C, D, E, G). Thus, the coreness of A  
 172 and F is 0. In that 1-core of the subgraph, nodes D and G have in-degree 1. While they are not part of

173 the 2-core of the graph, deleting them also renders B ineligible for 2-core. Thus, the coreness index of  
174 nodes B, D and G is 1. This process continues, until all nodes are assigned a coreness index. We call this  
175 coreness index, the **labor market centrality index** (LMCI) when applied to commuting networks.  
176 LMCI is a dimensionless number. The absolute scale of LMCI is not important, as the counties in the  
177 upper decile of the index are characterized as Strong Core counties. The counties above the third quartile  
178 and below the upper decile are categorized as Weak Core, and counties below the third quartile are  
179 Periphery. We then compare these categorizations to other conceptions of centrality and urbanity. To  
180 demonstrate the usefulness of this categorization, we use location quotients of different employment  
181 sectors and growth patterns to illustrate the differences. We use R (R Development Core Team 2017)  
182 and igraph package (Csardi and Nepusz 2006) for the analysis. We present the results using ggplot2  
183 (Wickham 2016) and tmap (Tennekes 2018) packages.

184

### 185 3.2. Data

186

187 We use the 2011–2015 county-to-county commuting flow data from the American Community Survey  
188 (2005). For the sake of exposition, we limit our analysis to the conterminous United States consisting of  
189 3,109 counties. In this American Community Survey (ACS) data, 134,869 pairs of counties have non-  
190 zero commuters, representing 1.4% of the possible links. The network is relatively sparse, a testament to  
191 the continuing importance of geographic distance for labor market integration. These links represent  
192 142.5 million commuters, of which 72% commuted within the same county. Using this data, we  
193 construct a directed network with self-loops, with counties as nodes. Two nodes (including the same  
194 node) are connected by an edge, if there are non-zero number of commuters from the residence county to  
195 workplace county. To simplify the computations, we scale the number of commuters logarithmically and  
196 use them as edge weights, a measure of strength of connection between the two counties.

197

### 198 3.3. Robustness Checks

199

200 Within county commuters account for a substantial portion of the commuting in the United States. To  
201 test the effect of within commuting on the regional structure, we removed them and repeated the process  
202 described in section 3.1. Furthermore, because ACS is a survey rather than a census, each commuting  
203 value has an associated Margin of Error (MOE). The 90% confidence interval MOE for the commuting  
204 between two counties ranged from 1 to 8,301. To account for the impact of MOE on the labor market



205 centrality index, we draw a random number from a normal distribution with mean as the estimate of the  
206 number of commuters and standard deviation as  $MOE/1.65$  for each pair of counties, truncating at 0. The  
207 distributions are assumed to be independent for pairs of counties. Using these random numbers as  
208 weights, we repeat the process (described in section 3.1) 1,000 times in a Monte Carlo simulation. To  
209 study the effect of the logarithmic transformation, we also use square root and linear transformation on  
210 the weights.

211

#### 212 4. Results

213

214 The results point to tightly connected large cores in the Northeastern United States that span Boston to  
215 Washington, D.C.; in Florida around Miami and Tampa; in Southern California around Los Angeles; and  
216 in Northern California around San Francisco (see Figure 2a). As can be expected, there are also  
217 numerous other smaller cores around Miami, Atlanta, Chicago, Detroit, Seattle, Denver and other cities.  
218 While it is tempting to conclude that core counties are counties with concentrated and large population,  
219 Figure 3 shows that at the upper tail of the distribution there is substantial variation in population. The  
220 empirical cumulative distribution function (ECDF) plots reveal interesting patterns (see Figure 2b). The  
221 central counties in MSAs clearly have high labor market centrality index compared to other types of  
222 counties. At the upper tail of the distribution, central MSA counties, the centrality index is not correlated  
223 with the population and is more reflective of the economic integration with the surrounding region (see  
224 Figure 3). The central  $\mu$ SA counties have much lower values than central MSA counties. However,  
225 some outlying MSA counties have higher index than central  $\mu$ SA (see Figure 2b).

226

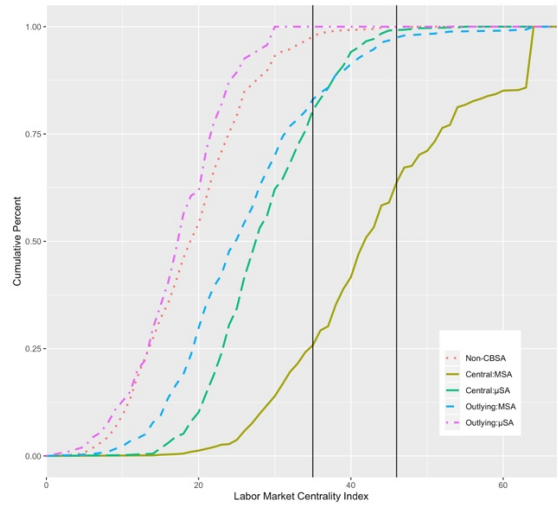
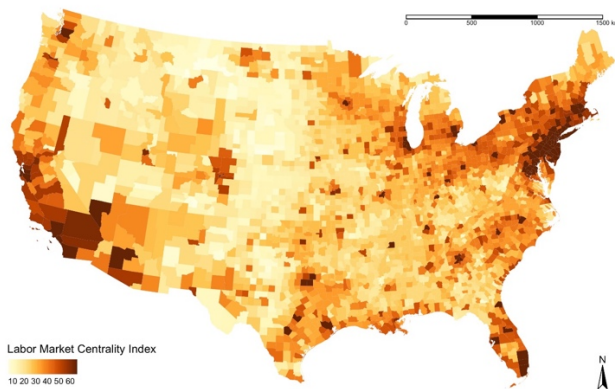
227 There are 280 counties classified as Strong Core and 471 classified as Weak Core. The rest are in the  
228 periphery (see Figure 3 & Figure 4). More importantly, many of the outlying  $\mu$ SA counties have lower  
229 index than non-CBSA counties (see Figure 2b). These unexpected results point to the need for closer  
230 examination of the classification that relies on node attributes. We need to rethink our understanding of  
231 the regional structure of the metropolitan USA and its relationship to the underlying labor market  
232 networks.

233

234 Figure 2 Distribution of the labor market centrality index based on the commuting network among counties in the conterminous United States  
235 a) Spatial distribution b) Empirical cumulative distribution by different types of counties defined by OMB.

236

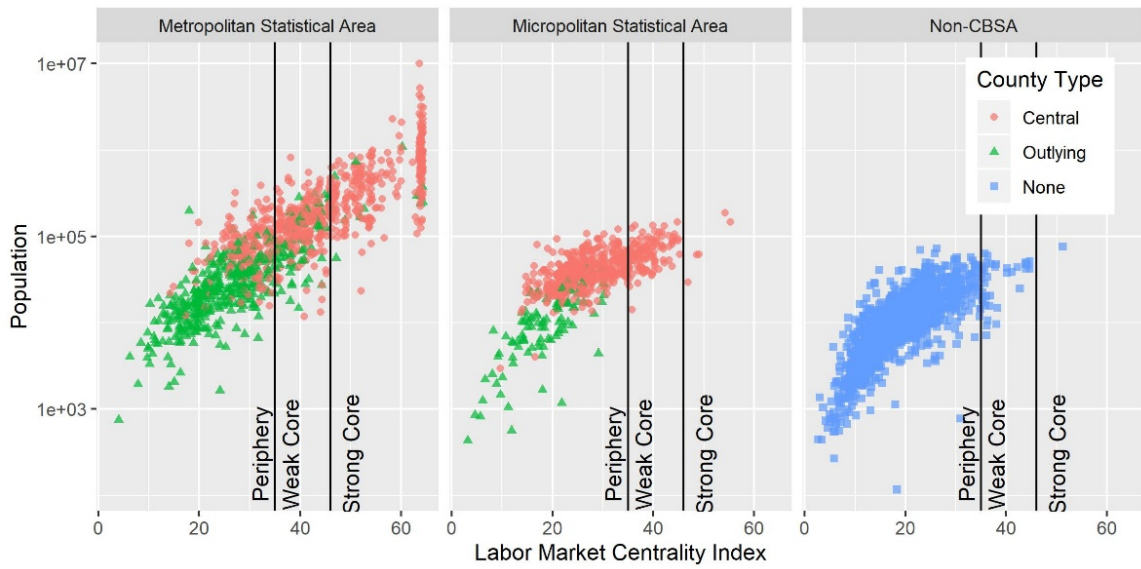
237



238

239  
240

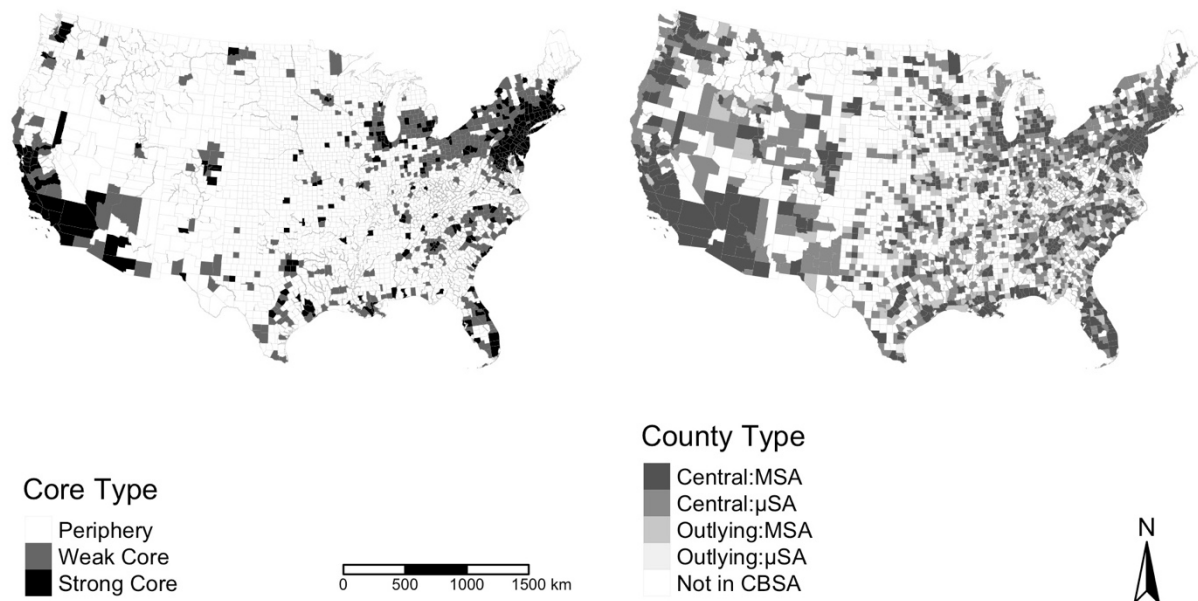
Figure 3 Labor market centrality index relative to population of the county. County type is by OMB. Strong and weak core are defined by labor market centrality index cut at quantiles 0.9, 0.75 respectively. Y-axis is logarithmically transformed for illustration.



241

242  
243  
244

Figure 4 Comparing the spatial distributions of core and periphery using different definitions a) Results from the network analysis. Strong and weak core are defined by labor market centrality index cut at quantiles 0.9, 0.75 respectively b) OMB categorizations of central and outlying counties in different CBSA types.



245  
246  
247  
248

#### 249 4.1. Strong and weak core and relationship to OMB categorization

250 While it is tempting to conclude that core counties are counties with concentrated and large population,  
 251 Figure 3 shows that at the upper tail of the distribution, there is substantial variation; i.e. counties in the  
 252 last index decile have population ranging from 23,000 (Fairfax city) to 10 million (Los Angeles County).  
 253 Thus, the upper tail of the labor market centrality index is not directly correlated with the population and  
 254 is more reflective of the economic integration with the surrounding region. Within MSAs, many of the  
 255 outlying counties (63) are part of the Weak core, but fewer are in the Strong core (11) (see Figure 3). Of  
 256 the 560 Central counties in μSAs, 104 are classified as Weak Core and 5 are classified as Strong Core.  
 257 These counties do not necessarily have large populations, but few of them are over 100,000 people (see  
 258 Figure 3). There is one county (Sullivan, New York) that is not part any CBSA but belongs to a strong  
 259 core. Similarly, two counties in New York, and one each in Connecticut, Pennsylvania and North Dakota  
 260 are part of μSAs, but are classified as Strong Core. Two such counties have more than a 100,000 people.  
 261 In contrast, there are no outlying counties in μSA that are a Strong or Weak Core. Six MSA counties  
 262 with less than 100,000 population are classified as Strong Core and 98 as Weak Core.

263

264 In total, 638 Central counties are not part of Strong or Weak core. While these counties have urban  
265 populations above the CBSA thresholds specified, they have fewer commuters both to other nodes as  
266 well as to themselves, implying a comparatively weak local and regional economy. In general, the  
267 population of these peripheral counties is lower than core counties. However, it is not universally true;  
268 65 counties (most of them MSA Central counties) have more than 100,000 people are in the periphery.  
269 These disagreements in classifications provide a productive starting point to analyze the role of ‘small’  
270 non-urban counties in the regional economy as well as large urban counties that are experiencing  
271 economic stagnation and decline.

272

273 There are 92 distinct geographical clusters of Strong Core counties (defined by queen contiguity), with  
274 the biggest one comprising of 109 counties stretching from Portland, Maine to Northern Virginia. The  
275 second biggest cluster is the 28-county collection in California, from San Diego to Santa Rosa. The rest  
276 of the geographic clusters are comprised of 1 to 7 counties, with 65% of them being a single county.  
277 With the inclusion of Weak Core counties, the number of geographic clusters to increases to 108: 23 of  
278 the clusters are a collection of Weak and Strong core counties; 49 of the clusters are only comprised of  
279 Weak Core counties.

280

#### 281 *4.2. Robustness checks and uncertainty estimates*

282 When self-loops were removed, the indices with and without them, for each county differ on average by  
283 3.9 with a maximum of 6 and a minimum of 2. The correlation coefficient between the indices with and  
284 without the loops is 0.99, implying that the main conclusions are not affected by the consideration of  
285 intra-county commuting. The categorizations of Weak and Strong Cores are not affected.

286

287 In the Monte Carlo simulations accounting for the MOE in the commuting (described in section 3.3), the  
288 indices of a county have a maximum range of 15 and minimum of 1, with an average of 6.7. The  
289 standard deviation, however, is small (<2.3). Counties with higher (though not large) variance in the  
290 labor market centrality index are relatively sparsely populated, are near economic centers and are more  
291 likely to be in the periphery, though there are some exceptions (e.g. Wake County in North Carolina and  
292 Duval County in Florida).

293

294 The precise monotonic transformation is largely irrelevant to main conclusions. While the main results  
295 are presented with log transformation of the number of commuters, we experimented with square root,

296 linear transformations and recovered the main results but for the variations in rounding to integers. The  
 297 categorization of Weak and Strong Cores is not affected since the cuts are based on quantiles. The  
 298 rounding to integers does not pose a major problem to the robustness of results as the estimates of  
 299 commuters come with margins of error and the rounding errors are subsumed within them.

300

## 301 5. Discussion

302 We find that micropolitan areas are almost comprised of exclusively periphery counties, but metropolitan  
 303 areas are have a wide diversity. Similarly, other measures of urbanity show that places outside of urban  
 304 areas are fairly consistently classified as Periphery but counties with large populations or dense  
 305 population are not necessarily Strong Core counties. This is because the Labor Market Centrality Index  
 306 captures economic aspects, mainly by using total commuting flows which is closely related to  
 307 employment. Strong Core counties tend to be those that have an expanding economy, regardless of  
 308 population size.

309

### 310 5.1. Regional Structure of the Statistical Areas

311

312 Of the 378 MSAs, only 38 contain all three types of counties and none of the  $\mu$ SA contains all three. In  
 313 a substantial number of cases, MSAs contain exclusively one type of county; for example, 56 MSAs are  
 314 exclusively strong core counties and 70 MSAs are exclusively comprised on Periphery counties (see  
 315 Table 2). Interestingly, five  $\mu$ SAs (Hudson, NY, Oneonta, NY, Pottsville, PA, Torrington, CT, Williston,  
 316 ND) comprise exclusively of Strong Core counties. 157 MSAs contained no counties that are Periphery  
 317 and 222 MSAs contained no Strong Core counties. 103  $\mu$ SA have no Periphery counties and 542 have no  
 318 Strong Core. As mentioned before one county that is not part of CBSA is considered a Strong Core, 29  
 319 of them are Weak Core; a vast majority (97%) of the non-CBSA counties are Peripheral counties.

320

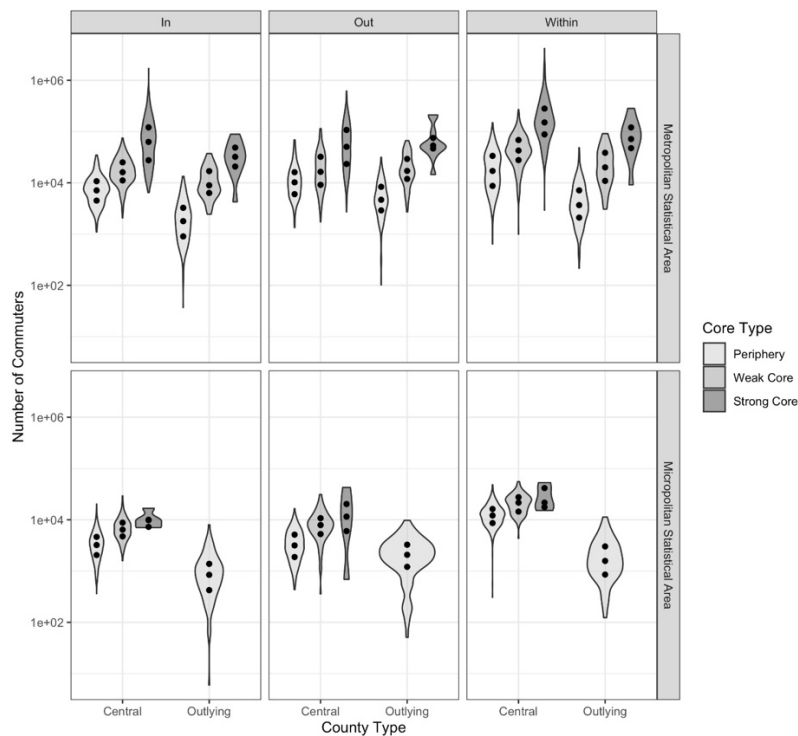
321 Table 2 Differences in the Structure of MSA and  $\mu$ SA based on Labor Market Centrality Index

	Number of MSA (378)				Number of $\mu$ SA (547)			
	Exclusive	Combination (AND)			Exclusive	Combination (AND)		
Periphery	70		77	88	439		5	0
Weak Core	75	26			98	0		
Strong Core	56				5			

322

323 Strong Core counties, on average, have higher number of both in, out and within commuters in both  
 324 MSA and  $\mu$ SA (see Figure 5). Within MSAs, there is marked difference in the in-commuters within  
 325 central counties that are classified as strong core or periphery, while such stark difference is absent  
 326 among the distributions of the commuters in the weak core and periphery counties. However, among the  
 327 outlying counties, there are substantial differences in the commuters in periphery, weak core and strong  
 328 core counties in all types of commuters, with a clear gradation. These differences are noticeably absent  
 329 in the  $\mu$ SAs.

330 Figure 5 Patterns of commuting in different types of counties within CBSAs. Violin plots represent the distributions while the interquartile range  
 331 and the median are represented by points. Y-axis is logarithmically transformed for illustration.



332

333

### 334 5.2. Relationship of Coreness to Urbanity

335

336 Very few (29) of the Strong Core counties have less than 250,000 people (see Table 3). However, the  
 337 existence of large population does not make the county part of the Core. In fact, there are as many  
 338 counties with million plus people that are in the Periphery as there are in the Strong and Weak Cores.  
 339 These are counties that are primarily residential counties that do not have strong economic attractors  
 340 within them. Small counties (less than 250,000 population) with metro areas are far more likely to be in  
 341 the Periphery (223) than in the Strong Core (23), though there are some exceptions. Counties such as

342 Litchfield, CT and Schuylkill, PA have less than 250,000 people and are part of the Strong Core. On the  
 343 other hand, Sullivan, NY, a Strong Core county with less than 80,000 population is not part of any  
 344 CBSA region. There are 6 counties that RUCC classifies as non-metro (categories 4 through 9) that  
 345 Labor Market Centrality Index identifies as Strong Core. The vast majority of nonmetro counties,  
 346 however, are classified as periphery.

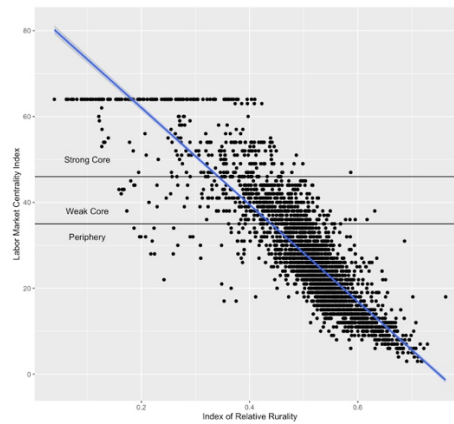
347 Table 3 Contingency table of different types of classification of counties based on urbanity with the Core-Periphery classification (source: ERS,  
 348 USDA 2013 and NCHS, CDC 2014)

County Type in Different Classification Schemes <sup>†</sup>	Periphery	Weak Core	Strong Core
<i>ERS Rural Urban Continuum Code Classification</i>			
Counties in metro areas of 1 million population or more	149	125	158
Counties in metro areas of 250,000 to 1 million population	187	95	93
Counties in metro areas of fewer than 250,000 population	212	118	23
Urban population of 20,000 or more, adjacent to a metro area	140	71	3
Urban population of 20,000 or more, not adjacent to a metro area	78	11	0
Urban population of 2,500 to 19,999, adjacent to a metro area	553	39	1
Urban population of 2,500 to 19,999, not adjacent to a metro area	414	9	2
Completely rural or less than 2,500 urban population, adjacent to a metro area	218	1	0
Completely rural or less than 2,500 urban population, not adjacent to a metro area	406	2	0
<i>NCHS Urban Rural Classification</i>			
Large Central Metro	0	0	68
Large Fringe Metro	149	128	91
Medium Metro	187	91	91
Small Metro	212	119	24
Micropolitan	528	104	5
Non-core	1281	29	1

349  
 350 Discretization might produce spurious relationships because of edge effects therefore it is useful to look  
 351 at the underlying continuous variables. Waldorf and Kim (2018) fashion a continuous Index of Relative  
 352 Rurality from 0 (most urban) to 1 (most rural) based on population size, density, remoteness and built-up  
 353 area of a county. The Spearman's correlation between IRR and the Labor Market Centrality Index is  
 354 -0.88, suggesting that the more urban the county, the more likely it is a Core county. However, closer  
 355 examination suggests that much of this correlation is driven by rural counties with low Labor Market  
 356 Centrality Index values. At the top end of the Labor Market Centrality Index and the lower end of the  
 357 IRR (i.e., urban), there is substantial variance (see Figure 6).

<sup>†</sup> The counts in the NCHS and ERS classification doesn't match up with OMB classifications due to different data vintages. For example, Garfield County, OK is part of Enid MSA according to OMB in 2015 but is part of Enid  $\mu$ SA according to NCHS and ERS, which relied on OMB 2013 delineations. Likewise, 18  $\mu$ SA counties in 2015 delineations and were classified as Non-core by NCHS and ERS.

358 Figure 6 Relationship between Labor Market Centrality Index and IRR. Source: Waldorf &amp; Kim (2018)



359

360 *5.3. Economic Specialization of Core and Periphery*

361

362 The previous measures indicate population size and density but the Labor Market Centrality Index  
 363 reflects economic dimensions of the labor market. In counties with population over 250,000 there is a  
 364 marked difference in the employment patterns. The median employment of Strong Core, Weak Core and  
 365 Peripheral counties is around 0.365, 0.171 and 0.148 million suggesting a strong economic  
 366 differentiation. More crucially the economic structure is also different among the different types of  
 367 counties. On average, Strong Core counties have proportionally more private employment with median  
 368 location quotient greater than 1 (see Table 4). Local government employment on the other hand is much  
 369 higher in the Periphery counties than in the Strong Core counties. Within Private sector employment,  
 370 Periphery counties as well as Weak Core, on average, specialize in goods producing industries, while the  
 371 Strong Core counties specialize in service providing industries. Places with expanding economies tend to  
 372 be more specialized in private employment instead of public employment, which can be thought of as  
 373 subsistence employment: jobs that enable people to continue to live there. The US economy has  
 374 expanded much more dramatically in the service sector over the past 70 years than in the goods-  
 375 producing industries (Buera & Kaboski 2012). Strong Core counties tend to be more specialized in  
 376 services while Weak Core and Periphery counties have proportionately more jobs in goods-producing  
 377 industries.

378

379

380



381 Table 4 Employment Location Quotient by County Type. Median values are displayed in the table with IQR in the parenthesis. Source: Bureau  
382 of Labor Statistics 2016.

County Type	Location Quotient					
	Federal Government	State Government	Local Government	Private		
				Goods producing	Service providing	
Periphery	0.53 (0.53)	0.61 (0.99)	1.74 (0.98)	1.42 (0.99)	0.75 (0.22)	
Weak Core	0.35 (0.58)	0.65 (0.89)	1.15 (0.48)	1.24 (0.88)	0.9 (0.19)	
Strong Core	0.62 (0.8)	0.67 (0.88)	0.9 (0.37)	0.85 (0.41)	1.02 (0.14)	

383

384 Table 5 Change in Employment by County Type. Median Values are displayed in the table with IQR in the parenthesis. Source: Bureau of  
385 Economic Analysis 2019

County Type	Employment Growth	
	2001-2008	2008-2015
Periphery	2.98% (12.73)	-0.52% (8.62)
Weak Core	7.67% (16.95)	2.56% (10.1)
Strong Core	7.52% (10.32)	5.05% (8.81)

386

387 It is illustrative to see changes in the employment pre and post-recession in different types of counties.  
388 While the Weak Core counties grew (in terms of number of jobs) roughly at the same rate as the Strong  
389 Core counties pre-recession (2001-2008), the recovery in the post-recession has been twice as strong in  
390 the Strong Core counties in the post-recession (see Table 5). The recovery seems to have bypassed the  
391 Periphery counties; while they grew at a healthy 3% before the recession, they contracted by 0.5% after  
392 the recession. In part, these numbers can be explained due the spatial sorting of specializations and the  
393 changing nature of the economy. However, these distinctions are not as stark, if we use the Central and  
394 Outlying distinctions of OMB. Central counties (on average) marginally grew faster compared to  
395 Outlying counties (1.6% vs. 0.63%) during the post-recession, even while they had similar growth rates  
396 pre-recession (6.75 vs. 6.45). However, Central counties with MSA significantly outpaced Central  
397 counties within  $\mu$ SA in post-recession recovery (4.3% vs. -0.84%). This, together with the specialization  
398 in service industries indicates that it is not the population size of the county that is related to the  
399 economy but rather its place in the regional network. We do not make any claims as to the causal  
400 relationship between the position in the network and the economic growth.

401

402

#### 403 5.4. *Limitations and Future Work*

404

405 County as unit of analysis in a commuting network may be useful for policy purposes, but micro level  
406 regional structures of the commuting network can also be inferred using a finer geographic scale such as  
407 census tracts. However, at a census tract level, the MOE is substantial and Monte Carlo simulations take  
408 significant computational resources. Nevertheless, future work should understand the role of geographic  
409 scale in determining the centrality of places. Core periphery structure in tract based commuting network  
410 can be used to extend and refine the work of He et. al (2020), who characterize the overlapping  
411 communities in a commuting network and distinguish between nodal and non-nodal clusters of counties.  
412 Such work is also a natural extension of Hartley et. al (2016) who identify the employment centers in a  
413 tract based commuting network. It would useful to identify if these centers identified using McMillen  
414 (2001) correspond closely with the identified cores.

415

416 The labor market centrality index has been calculated from ACS data that is primarily cross-sectional.  
417 Longer term time trends in commuting patterns might be useful to more fully characterize the  
418 positionality of a node in a network. Other networks such as business transactions can supplement the  
419 information in commuting networks to get a more complete understanding of the place in the regional  
420 hierarchies.

421

#### 422 **6. Conclusion**

423

424 There are many ways to understand human settlements. In this paper, we looked at the regional structure  
425 from a network perspective. We found that how a county functions within the network of human  
426 settlement across the continental US is based on population and economic activity. Our typology reflects  
427 economic dimensions in addition to population and density.

428

429 Metropolitan regions are formed around economic activity and therefore reflect economic centers but  
430 existing typologies do not characterize the strength and nature of the regional economy well. Focusing  
431 on the role of the county in the network through commute patterns illuminates not just how central a  
432 county is in the labor market but also broadly demonstrates the strength of the economy. This is  
433 independent of the size of the population. Although there is some relationship between the size of the

434 population and the size of the economy, there were some small counties with a lot of commute flows and  
435 large counties that had very little commuting.

436

437 Basing the index on total commuting rather than the number of commuters relative to the population of  
438 the sending county, more closely mimics jobs and therefore reflects the character of the economy. Places  
439 with large populations but with little economic activity to have a lower index. Rural places that may have  
440 a lot of internal commuting but not a lot of commuting from neighboring areas will also have a lower  
441 score. Places that have a lot of economic activity relative to the population size will score high because  
442 they not only have a lot of internal commuting but also a lot of in-commuting from surrounding counties.

443

444 This typology has the potential to be more dynamic than the OMB definition of metropolitan and  
445 micropolitan areas. Every time the OMB definitions are updated, there are changes to the delineations of  
446 Metropolitan and micropolitan areas but those changes are mostly additive. Even regions in economic  
447 depressions do not lose their metropolitan status because the overall population continues to grow. By  
448 basing the typology on total commute flows, it reflects a region's total economic activity and the  
449 connection between residents and that economic activity. In addition, our index categorizes counties into  
450 Strong Core, Weak Core, and Periphery based on the score relative to other counties.

451

452 Categorizing counties based on their function in the network of human settlements is a useful way to  
453 understand the integration of population and economy. It shares some similarities with other typologies  
454 focused on commuting flows. However, it has the unique feature of reflecting the economic strength of  
455 the region in a more dynamic way than other categorizations and indices.

456

457

458 **References**

- 459 Agryzkov, Taras, Leandro Tortosa, José F Vicent, and Richard Wilson. 2019. 'A Centrality Measure for Urban  
460 Networks Based on the Eigenvector Centrality Concept'. *Environment and Planning B: Urban Analytics  
461 and City Science* 46(4): 668–89.
- 462 Batagelj, Vladimir, and Matjaž Zaveršnik. 2011. 'Fast Algorithms for Determining (Generalized) Core Groups in  
463 Social Networks'. *Advances in Data Analysis and Classification* 5(2): 129–45.
- 464 Bonacich, Phillip. 1972. 'Factoring and Weighting Approaches to Status Scores and Clique Identification'. *Journal  
465 of mathematical sociology* 2(1): 113–20.
- 466 Cingolani, Isabella, Pietro Panzarasa, and Lucia Tajoli. 2017. 'Countries' Positions in the International Global  
467 Value Networks: Centrality and Economic Performance'. *Applied Network Science* 2(1): 21.
- 468 Congressional Research Service. 2014. *Metropolitan Area Designations by OMB: History, 2010 Standards, and  
469 Uses*. Washington DC. CRS Report.  
470 [https://www.everycrsreport.com/files/20140606\\_R42005\\_cc88d5c754b797d095e0880142d7c28aa739d87  
471 1.pdf](https://www.everycrsreport.com/files/20140606_R42005_cc88d5c754b797d095e0880142d7c28aa739d871.pdf).
- 472 Cook, Peggy J., and Karen Mizer. 1994. *ERS Typology Revised and Updated*. Washington DC: US Department of  
473 Agriculture. Rural Development Research Report.  
474 <https://naldc.nal.usda.gov/download/CAT10829253/PDF> (July 4, 2019).
- 475 Csardi, Gabor, and Tamas Nepusz. 2006. 'The Igraph Software Package for Complex Network Research'.  
476 *InterJournal Complex Systems*: 1695.
- 477 Cutter, Susan L., Kevin D. Ash, and Christopher T. Emrich. 2016. 'Urban–Rural Differences in Disaster  
478 Resilience'. *Annals of the American Association of Geographers*.  
479 <https://www.tandfonline.com/doi/abs/10.1080/24694452.2016.1194740> (November 19, 2019).
- 480 De Lew, Nancy. 1992. 'Geographic Classification of Hospitals: Alternative Labor Market Areas'. *Health Care  
481 Financing Review* 14(2): 49–58.
- 482 Degenne, Alain, and Michel Frose. 1999. *Introducing Social Networks*. 1 edition. London ; Thousand Oaks: SAGE  
483 Publications Ltd.
- 484 Derrible, Sybil. 2012. 'Network Centrality of Metro Systems'. *PLOS ONE* 7(7): e40575.
- 485 Dorogovtsev, S. N., A. V. Goltsev, and J. F. F. Mendes. 2006. 'K-Core Organization of Complex Networks'.  
486 *Physical Review Letters* 96(4). <http://arxiv.org/abs/cond-mat/0509102> (March 18, 2019).
- 487 Ducruet, César, and Laurent Beauguitte. 2014. 'Spatial Science and Network Science: Review and Outcomes of a  
488 Complex Relationship'. *Networks and Spatial Economics* 14(3): 297–316.
- 489 Eidsaa, Marius, and Eivind Almaas. 2013. '\$s\$-Core Network Decomposition: A Generalization of \$k\$-Core  
490 Analysis to Weighted Networks'. *Physical Review E* 88(6): 062819.
- 491 ERS. 2013. 'Rural-Urban Continuum Codes'. [https://www.ers.usda.gov/data-products/rural-urban-continuum-  
492 codes/documentation/](https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/) (November 18, 2019).
- 493 Farrigan, Tracy. 2019. 'Rural Employment and Unemployment'. [https://www.ers.usda.gov/topics/rural-economy-  
494 population/employment-education/rural-employment-and-unemployment/](https://www.ers.usda.gov/topics/rural-economy-population/employment-education/rural-employment-and-unemployment/) (April 27, 2020).

- 495 Fowler, Christopher, S., Leif Jensen, and Danielle Rhubart. 2018. 'Assessing U.S. Labor Market Delineations for  
496 Containment, Economic Core, and Wage Correlation.'
- 497 Freeman, Linton C. 1978. 'Centrality in Social Networks Conceptual Clarification'. *Social networks* 1(3): 215–39.
- 498 Han, Yicheol, and Stephan J. Goetz. 2015. 'The Economic Resilience of U.S. Counties during the Great  
499 Recession'. *Review of Regional Studies* 45(2): 131–49.
- 500 Han, Yicheol, and Stephan J. Goetz. 2019. 'Overlapping Labour Market Areas Based on Link Communities'.  
501 *Papers in Regional Science* 98(1): 539–53.
- 502 Hartley, D. A., N. Kaza, and T. William Lester. 2016. 'Are America's Inner Cities Competitive? Evidence From  
503 the 2000s'. *Economic Development Quarterly* 30(2): 137–58.
- 504 He, Mark et al. 2020. 'Demarcating Geographic Regions Using Community Detection in Commuting Networks'.  
505 *PLOS ONE* 15(4): e0230941.
- 506 Hewitt, Maria. 1989. *Defining 'Rural' Areas: Impact on Health Care Policy and Research*. Washington DC: US  
507 Congress. Staff Paper. [http://govinfo.library.unt.edu/ota/Ota\\_2/DATA/1989/8912.PDF](http://govinfo.library.unt.edu/ota/Ota_2/DATA/1989/8912.PDF).
- 508 Ingram, D. D., and S. J. Franco. 2012. 'NCHS Urban-Rural Classification Scheme for Counties.' *Vital and health  
509 statistics. Series 2, Data evaluation and methods research* (154): 1–65.
- 510 Isserman, A.M. 2005. 'In the National Interest: Defining Rural and Urban Correctly in Research and Public  
511 Policy'. *International Regional Science Review* 28(4): 465–99.
- 512 Kirkley, Alec, Hugo Barbosa, Marc Barthelemy, and Gourab Ghoshal. 2018. 'From the Betweenness Centrality in  
513 Street Networks to Structural Invariants in Random Planar Graphs'. *Nature Communications* 9(1): 1–12.
- 514 Klove, Robert C. 1952. 'The Definition of Standard Metropolitan Areas'. *Economic Geography* 28(2): 95–104.
- 515 Lipscomb, C. A., and R.V. Kashbrasiev. 2008. 'County Typologies to Inform Job Tax Credit Policy in Georgia -  
516 ProQuest'. *Review of Regional Studies* 38(2): 233–50.
- 517 Maggioni, Mario A., and Teodora Erika Uberti. 2009. 'Knowledge Networks across Europe: Which Distance  
518 Matters?' *The Annals of Regional Science* 43(3): 691–720.
- 519 McMillen, Daniel P. 2001. 'Nonparametric Employment Subcenter Identification'. *Journal of Urban Economics*  
520 50(3): 448–73.
- 521 Narang, Vipin et al. 2015. 'Automated Identification of Core Regulatory Genes in Human Gene Regulatory  
522 Networks'. *PLOS Computational Biology* 11(9): e1004504.
- 523 NCHS. 2014. *2013 NCHS Urban-Rural Classification Scheme for Counties*. Hyattsville, MD: Centers for Disease  
524 Control, U.S. Department of Health and Human Services.  
525 [https://www.cdc.gov/nchs/data/series/sr\\_02/sr02\\_166.pdf](https://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf) (March 2, 2018).
- 526 Neal, Zachary. 2011. 'Differentiating Centrality and Power in the World City Network'. *Urban Studies* 48: 2733–  
527 48.
- 528 Nelson, Garrett Dash, and Alasdair Rae. 2016. 'An Economic Geography of the United States: From Commutes to  
529 Megaregions'. *PLOS ONE* 11(11): e0166083.
- 530 Newman, M. E. J. 2016. 'Mathematics of Networks'. In *The New Palgrave Dictionary of Economics*, London:  
531 Palgrave Macmillan UK, 1–8. [https://doi.org/10.1057/978-1-349-95121-5\\_2565-1](https://doi.org/10.1057/978-1-349-95121-5_2565-1) (November 19, 2019).

- 532 Nystuen, John D., and Michael F. Dacey. 1961. 'A Graph Theory Interpretation of Nodal Regions'. *Papers of the*  
533 *Regional Science Association* 7(1): 29–42.
- 534 R Development Core Team. 2017. 'R: A Language and Environment for Statistical Computing'. Vienna, Austria.  
535 <http://www.R-project.org>.
- 536 Scala, Dante J., and Kenneth M. Johnson. 2017. 'Political Polarization along the Rural-Urban Continuum? The  
537 Geography of the Presidential Vote, 2000–2016'. *The ANNALS of the American Academy of Political and*  
538 *Social Science*. <http://journals.sagepub.com/doi/full/10.1177/0002716217712696> (November 19, 2019).
- 539 Seidman, Stephen B. 1983. 'Network Structure and Minimum Degree'. *Social Networks* 5(3): 269–87.
- 540 Sigler, Thomas, and Kirsten Martinus. 2016. 'Extending beyond "world Cities" in World City Network (WCN)  
541 Research: Urban Positionality and Economic Linkages through the Australia-Based Corporate Network'.  
542 *Environment and Planning A* 49.
- 543 Tennekes, Martijn. 2018. 'tmap: Thematic Maps in R'. *Journal of Statistical Software* 84(6): 1–39.
- 544 Tolbert, Charles M., and Molly Sizer. 1996. *US Commuting Zones and Labor Market Areas: A 1990 Update*.  
545 Washington, DC: Rural Economy Division, Economic Research Service, US Department of Agriculture.  
546 <https://usa.ipums.org/usa/resources/volii/cmz90.pdf> (January 30, 2014).
- 547 Tong, Daoqin, and David A. Plane. 2014. 'A New Spatial Optimization Perspective on the Delineation of  
548 Metropolitan and Micropolitan Statistical Areas'. *Geographical Analysis* 46(3): 230–49.
- 549 US Census Bureau. 2018. '2011-2015 5-Year American Community Survey Commuting Flows'.  
550 <https://www.census.gov/data/tables/2015/demo/metro-micro/commuting-flows-2015.html> (March 17,  
551 2019).
- 552 Verma, T., N. a. M. Araújo, and H. J. Herrmann. 2014. 'Revealing the Structure of the World Airline Network'.  
553 *Scientific Reports* 4: 5638.
- 554 Waldorf, Brigitte, and Ayoung Kim. 2018. 'The Index of Relative Rurality (IRR): US County Data for 2000 and  
555 2010'. <https://purr.purdue.edu/publications/2960/1>.
- 556 Wang, Man, Rachel Garshick Kleit, Jane Cover, and Christopher S. Fowler. 2012. 'Spatial Variations in US  
557 Poverty: Beyond Metropolitan and Non-Metropolitan'. *Urban Studies* 49(3): 563–85.
- 558 Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.  
559 <https://ggplot2.tidyverse.org>.
- 560  
561  
562  
563