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

Katherine Nesse Ph.D.
City of Bellevue, WA 98004
knesse@bellevuewa.org
Characterizing the Regional Structure in the United States: A county-based analysis of labor market centrality

Abstract

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.
1. Introduction

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

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

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

<table>
<thead>
<tr>
<th>County Type</th>
<th>CBSA Type</th>
<th>Central</th>
<th>Outlying</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSA</td>
<td>725</td>
<td>436</td>
<td>1,161</td>
</tr>
<tr>
<td></td>
<td>µSA</td>
<td>560</td>
<td>94</td>
<td>654</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1,285</td>
<td>530</td>
<td>1,815</td>
</tr>
</tbody>
</table>

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

There is often a conflation of urbanicity with metropolitan areas. Isserman (2005) identified the differences between the US Census demarcation of urban and rural and OMB delineation of metropolitan and non-metropolitan areas, even when they are frequently used interchangeably. The former is about separation of densely built from sparsely built places, while the latter is primarily about integration of residence and place-of-work. Consequently, OMB’s metropolitan areas include large swaths of rural lands, centered on urban counties. Nonetheless, the context of the metropolitan area becomes important feature for determination of the urbanity of a place. For example, the Economic Research Service (ERS) of the US Department of Agriculture created Rural-Urban Continuum Codes (RUCC) to distinguish counties based on the population size of the metro area and proximity to metro areas for a total of nine categories (3 metro and 6 non-metro)(ERS 2013). ERS has created a commuting rubric at the sub-county level as well. The Rural-Urban Community Area (RUCA) codes identify census tracts as metropolitan, micropolitan or nonmetropolitan and breaks these down based on the size of the commute flow (ERS 2019). This 10-category classification system is more detailed both geographically and categorically but is still based in the OMB system. The National Center for Health Statistics (NCHS) created their own urban-rural typology of counties (4 metro and 2 non-metro) skewed towards metropolitan areas, by arguing that they stand in for urban and rural distinctions (NCHS 2014). These categories also include the distinctions between central and fringe counties in large metro regions as well as based on the size of the region (large, medium, small). For example, Han and Goetz (2015) argue that resilience patterns are different across different types of counties. Using the economic recession of 2008 and the recovery pattern as evidenced by employment changes, they argue that counties with RUCC codes 1-5 (large population metro and non-metro areas) have more resistance rather than resilience. Interestingly, an USDA report conflates rural with non-metropolitan areas and argues that rural areas are slow to recover post-recession (Farrigan 2019). These concepts of centrality, urbanity and proximity, condition our understanding of disparities and challenges faced by different regions (e.g. Cutter, Ash, and Emrich...
While these are important, we argue that they are complementary and not interchangeable with centrality. In this paper, we seek to establish a different notion of centrality based on the position of the county in a commuting network. We aim to uncover the regional structure by explicitly focusing on the network rather than node attributes (unlike OMB’s characterization of centrality). In particular, counties (or equivalents) that have small population (either because of constricted boundaries or because of a lack of residential lands) might still be destination areas for commuting. On the other hand, counties that have large populations may not necessarily have sufficient economic activity to justify the Central designation. By focusing on the relationships among counties as a network, rather than the county attributes, we can begin to uncover some regional structures such as integration and core-periphery structures. The regions may be (multi) core based, or comprise exclusively of peripheral nodes. Our aim is to demonstrate that such understanding complements our conventional understanding of centrality used in the literature and policy analysis. We also seek to demonstrate meso-level regional structure rather than intra-regional spatial structure using a national analysis. We demonstrate that the positionality-based (as opposed to attribute based) centrality is correlated with economic growth and sectoral specialization. We also demonstrate that centrality is not correlated, in particular, with urbanity or population levels.

2. Network Analyses

Network analyses in regional science have a long history, though there have been significant divergences between geographers and network scientists (for an extensive literature review, see Ducruet and Beauguitte 2014). For example, Nystuen and Dacey (1961) use graph theory and commuting flows to identify regional hierarchies and nodal regions. Tong and Plane (2014) use spatial optimization techniques on the commuting network of all inter-county commuting linkages to identify clusters that rival OMB’s Metropolitan area delineations. Nelson and Rae (2016) use community detection techniques to derive the mega-region structures. Using different community detection techniques and statistical inference, He et al. (2020) propose that there are multiple overlapping regions in the United States, hitherto unrecognized by the OMB or the other delineations such as megaregions. These are but a few examples of use of graph theoretical analyses applied to regional science problems. Many of them rely on clustering of sub-geographies to construct a larger geographical region. Among network science
approaches in regional science, to our knowledge, very few have focused on centrality (see e.g. Neal 2011; Sigler and Martinus 2016).

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

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

3.1. s-core decomposition of a graph

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

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

$$s_{i} \equiv \sum_{j \in N_{i}^{-}} w_{ij}$$

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

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

The $s$–core decomposition is illustrated in Figure 1 for a directed graph with multiple edges including loops. The entire graph in the figure is part of 0-core. Nodes A and F have in-degree 0, and therefore are not part of the 1-core of the graph (subgraph induced by nodes B, C, D, E, G). Thus, the coreness of A 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...
the 2-core of the graph, deleting them also renders B ineligible for 2-core. Thus, the coreness index of nodes B, D and G is 1. This process continues, until all nodes are assigned a coreness index. We call this coreness index, the **labor market centrality index** (LMCI) when applied to commuting networks.

LMCI is a dimensionless number. The absolute scale of LMCI is not important, as the counties in the upper decile of the index are characterized as Strong Core counties. The counties above the third quartile and below the upper decile are categorized as Weak Core, and counties below the third quartile are Periphery. We then compare these categorizations to other conceptions of centrality and urbanity. To demonstrate the usefulness of this categorization, we use location quotients of different employment sectors and growth patterns to illustrate the differences. We use R (R Development Core Team 2017) and igraph package (Csardi and Nepusz 2006) for the analysis. We present the results using ggplot2 (Wickham 2016) and tmap (Tennekes 2018) packages.

### 3.2. Data

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

### 3.3. Robustness Checks

Within county commuters account for a substantial portion of the commuting in the United States. To test the effect of within commuting on the regional structure, we removed them and repeated the process described in section 3.1. Furthermore, because ACS is a survey rather than a census, each commuting value has an associated Margin of Error (MOE). The 90% confidence interval MOE for the commuting between two counties ranged from 1 to 8,301. To account for the impact of MOE on the labor market...
centrality index, we draw a random number from a normal distribution with mean as the estimate of the number of commuters and standard deviation as MOE/1.65 for each pair of counties, truncating at 0. The distributions are assumed to be independent for pairs of counties. Using these random numbers as weights, we repeat the process (described in section 3.1) 1,000 times in a Monte Carlo simulation. To study the effect of the logarithmic transformation, we also use square root and linear transformation on the weights.

4. Results

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

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

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

4.1. Strong and weak core and relationship to OMB categorization

While it is tempting to conclude that core counties are counties with concentrated and large population, Figure 3 shows that at the upper tail of the distribution, there is substantial variation; i.e. counties in the last index decile have population ranging from 23,000 (Fairfax city) to 10 million (Los Angeles County). Thus, the upper tail of the labor market centrality index is not directly correlated with the population and is more reflective of the economic integration with the surrounding region. Within MSAs, many of the outlying counties (63) are part of the Weak core, but fewer are in the Strong core (11) (see Figure 3). Of the 560 Central counties in μSAs, 104 are classified as Weak Core and 5 are classified as Strong Core. These counties do not necessarily have large populations, but few of them are over 100,000 people (see Figure 3). There is one county (Sullivan, New York) that is not part any CBSA but belongs to a strong core. Similarly, two counties in New York, and one each in Connecticut, Pennsylvania and North Dakota are part of μSAs, but are classified as Strong Core. Two such counties have more than a 100,000 people. In contrast, there are no outlying counties in μSA that are a Strong or Weak Core. Six MSA counties with less than 100,000 population are classified as Strong Core and 98 as Weak Core.
In total, 638 Central counties are not part of Strong or Weak core. While these counties have urban populations above the CBSA thresholds specified, they have fewer commuters both to other nodes as well as to themselves, implying a comparatively weak local and regional economy. In general, the population of these peripheral counties is lower than core counties. However, it is not universally true; 65 counties (most of them MSA Central counties) have more than 100,000 people in the periphery. These disagreements in classifications provide a productive starting point to analyze the role of ‘small’ non-urban counties in the regional economy as well as large urban counties that are experiencing economic stagnation and decline.

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

4.2. Robustness checks and uncertainty estimates

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

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

The precise monotonic transformation is largely irrelevant to main conclusions. While the main results are presented with log transformation of the number of commuters, we experimented with square root,
linear transformations and recovered the main results but for the variations in rounding to integers. The categorization of Weak and Strong Cores is not affected since the cuts are based on quantiles. The rounding to integers does not pose a major problem to the robustness of results as the estimates of commuters come with margins of error and the rounding errors are subsumed within them.

5. Discussion

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

5.1. Regional Structure of the Statistical Areas

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

Table 2 Differences in the Structure of MSA and µSA based on Labor Market Centrality Index

<table>
<thead>
<tr>
<th></th>
<th>Number of MSA (378)</th>
<th>Number of µSA (547)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exclusive Combination (AND)</td>
<td>Exclusive Combination (AND)</td>
</tr>
<tr>
<td>Periphery</td>
<td>70 77 88</td>
<td>439 98 0</td>
</tr>
<tr>
<td>Weak Core</td>
<td>75 26</td>
<td>5 0</td>
</tr>
<tr>
<td>Strong Core</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>
Strong Core counties, on average, have higher number of both in, out and within commuters in both MSA and μSA (see Figure 5). Within MSAs, there is marked difference in the in-commuters within central counties that are classified as strong core or periphery, while such stark difference is absent among the distributions of the commuters in the weak core and periphery counties. However, among the outlying counties, there are substantial differences in the commuters in periphery, weak core and strong core counties in all types of commuters, with a clear gradation. These differences are noticeably absent in the μSAs.

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

5.2. Relationship of Coreness to Urbanity

Very few (29) of the Strong Core counties have less than 250,000 people (see Table 3). However, the existence of large population does not make the county part of the Core. In fact, there are as many counties with million plus people that are in the Periphery as there are in the Strong and Weak Cores. These are counties that are primarily residential counties that do not have strong economic attractors within them. Small counties (less than 250,000 population) with metro areas are far more likely to be in the Periphery (223) than in the Strong Core (23), though there are some exceptions. Counties such as...
Litchfield, CT and Schuylkill, PA have less than 250,000 people and are part of the Strong Core. On the other hand, Sullivan, NY, a Strong Core county with less than 80,000 population is not part of any CBSA region. There are 6 counties that RUCC classifies as non-metro (categories 4 through 9) that Labor Market Centrality Index identifies as Strong Core. The vast majority of nonmetro counties, however, are classified as periphery.

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

<table>
<thead>
<tr>
<th>County Type in Different Classification Schemes†</th>
<th>Periphery</th>
<th>Weak Core</th>
<th>Strong Core</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ERS Rural Urban Continuum Code Classification</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counties in metro areas of 1 million population or more</td>
<td>149</td>
<td>125</td>
<td>158</td>
</tr>
<tr>
<td>Counties in metro areas of 250,000 to 1 million population</td>
<td>187</td>
<td>95</td>
<td>93</td>
</tr>
<tr>
<td>Counties in metro areas of fewer than 250,000 population</td>
<td>212</td>
<td>118</td>
<td>23</td>
</tr>
<tr>
<td>Urban population of 20,000 or more, adjacent to a metro area</td>
<td>140</td>
<td>71</td>
<td>3</td>
</tr>
<tr>
<td>Urban population of 20,000 or more, not adjacent to a metro area</td>
<td>78</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Urban population of 2,500 to 19,999, adjacent to a metro area</td>
<td>553</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>Urban population of 2,500 to 19,999, not adjacent to a metro area</td>
<td>414</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Completely rural or less than 2,500 urban population, adjacent to a metro area</td>
<td>218</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Completely rural or less than 2,500 urban population, not adjacent to a metro area</td>
<td>406</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>NCHS Urban Rural Classification</strong></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Large Central Metro</td>
<td>0</td>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>Large Fringe Metro</td>
<td>149</td>
<td>128</td>
<td>91</td>
</tr>
<tr>
<td>Medium Metro</td>
<td>187</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>Small Metro</td>
<td>212</td>
<td>119</td>
<td>24</td>
</tr>
<tr>
<td>Micropolitan</td>
<td>528</td>
<td>104</td>
<td>5</td>
</tr>
<tr>
<td>Non-core</td>
<td>1281</td>
<td>29</td>
<td>1</td>
</tr>
</tbody>
</table>

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

† 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 μSA according to NCHS and ERS, which relied on OMB 2013 delineations. Likewise, 18 μSA counties in 2015 delineations and were classified as Non-core by NCHS and ERS.
5.3. Economic Specialization of Core and Periphery

The previous measures indicate population size and density but the Labor Market Centrality Index reflects economic dimensions of the labor market. In counties with population over 250,000 there is a marked difference in the employment patterns. The median employment of Strong Core, Weak Core and Peripheral counties is around 0.365, 0.171 and 0.148 million suggesting a strong economic differentiation. More crucially the economic structure is also different among the different types of counties. On average, Strong Core counties have proportionally more private employment with median location quotient greater than 1 (see Table 4). Local government employment on the other hand is much higher in the Periphery counties than in the Strong Core counties. Within Private sector employment, Periphery counties as well as Weak Core, on average, specialize in goods producing industries, while the Strong Core counties specialize in service providing industries. Places with expanding economies tend to be more specialized in private employment instead of public employment, which can be thought of as subsistence employment: jobs that enable people to continue to live there. The US economy has expanded much more dramatically in the service sector over the past 70 years than in the goods-producing industries (Buera & Kaboski 2012). Strong Core counties tend to be more specialized in services while Weak Core and Periphery counties have proportionately more jobs in goods-producing industries.
Table 4 Employment Location Quotient by County Type. Median values are displayed in the table with IQR in the parenthesis. Source: Bureau of Labor Statistics 2016.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Periphery</td>
<td>0.53 (0.55)</td>
<td>0.61 (0.99)</td>
<td>1.74 (0.98)</td>
<td>1.42 (0.99)</td>
<td>0.75 (0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak Core</td>
<td>0.35 (0.58)</td>
<td>0.65 (0.89)</td>
<td>1.15 (0.48)</td>
<td>1.24 (0.88)</td>
<td>0.9 (0.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong Core</td>
<td>0.62 (0.8)</td>
<td>0.67 (0.88)</td>
<td>0.9 (0.37)</td>
<td>0.85 (0.41)</td>
<td>1.02 (0.14)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>County Type</th>
<th>Employment Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periphery</td>
<td>2.98%</td>
</tr>
<tr>
<td></td>
<td>(12.73)</td>
</tr>
<tr>
<td>Weak Core</td>
<td>7.67%</td>
</tr>
<tr>
<td></td>
<td>(16.95)</td>
</tr>
<tr>
<td>Strong Core</td>
<td>7.52%</td>
</tr>
<tr>
<td></td>
<td>(10.32)</td>
</tr>
</tbody>
</table>

It is illustrative to see changes in the employment pre and post-recession in different types of counties.

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

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

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

6. Conclusion

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

Metropolitan regions are formed around economic activity and therefore reflect economic centers but existing typologies do not characterize the strength and nature of the regional economy well. Focusing on the role of the county in the network through commute patterns illuminates not just how central a county is in the labor market but also broadly demonstrates the strength of the economy. This is independent of the size of the population. Although there is some relationship between the size of the
population and the size of the economy, there were some small counties with a lot of commute flows and large counties that had very little commuting.

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

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

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


Han, Yicheol, and Stephan J. Goetz. 2019. ‘Overlapping Labour Market Areas Based on Link Communities’. Papers in Regional Science 98(1): 539–53.


https://purr.purdue.edu/publications/2960/1.


https://purr.purdue.edu/publications/2960/1.