Regional Economic Tightness from Rural to Populous Regions

Keith Waters[1]*, Shade T. Shutters[2]

1 Schar School of Policy and Government, George Mason University, Fairfax, Virginia, USA
2 School of Complex Adaptive Systems, Arizona State University, Tempe, Arizona, USA
* Corresponding author: kwaters2@gmu.edu

Abstract

Regional economies are characterized by networks of interactions between individual elements and are thus quintessential complex systems. Analyzing the relatedness of various aspects of the regional economies, such as exports, industries, occupations, and technologies, using methods from complexity science is becoming commonplace. However, current work has focused nearly exclusively on regional economic complexity of more urbanized regions within countries, if not entire countries themselves. Less populous regions are typically over-looked and rural regions are almost entirely absent from discussions. This paper seeks to fill this gap by examining less populous and rural regions from a complexity economics perspective. Using a previously developed metric of economic connectivity, called tightness, we examine skills space and industry structure of metropolitan, micropolitan, and rural regions in the United States. We find that measures of economic complexity are contingent on method of spatial aggregation. We further show that while the least and the most populous regions have the highest tightness, the composition of skills in the least and most populous regions differs markedly.

Introduction

Analysis of regional economies from the basis that they are complex systems has surged in recent years. Complex systems are characterized by independent, yet networked, actors whose interactions result in system-level dynamics that are more than the simple sum of the parts (Slaper, 2019). Regional economies are quintessential examples of complex systems; regional economies are non-linear summations of the interactions of economic agents via their economic networks. Despite all of the work done on regional economies as complex systems, there has been a near unanimous focus on more populated regional economies. The focus on more populous regions is likely due to data limitations in rural regions as well as the supra-linear nature of economic activity, which provides results such as output per capita increasing faster than population, which is well-documented in the urban scaling literature (West, 2017). Such results, however, have perhaps distracted from the reality that less populated and rural regions are nonetheless complex systems that are also inherently of interest. Indeed, less populous and rural regions are worthy of more attention given structural disadvantages resulting from size documented in the scaling literature.

In this paper, we analyze regional economies as complex adaptive systems with an emphasis on less populated regions. Using a variety of geographic units that cover the entire U.S., we compare outcomes in areas spanning a wide range of population levels. The goal of this paper is twofold. First, we seek to make theoretical contributions to rural regions in both the regional economics and rural studies literature. Second, we seek to identify economic attributes important for regional economic policy making, particularly in less populated and rural regions.

This study makes several contributions. First, it systematically assesses how choice of areal units affects both industry co-occurrence and economic tightness, two measures frequently used to examine regional economies as complex systems. Second, it examines the impact of regional economic tightness on regional measures of regional economic performance before controlling for population. Finally, it analyzes how population impacts a region's location within the skills-space, previously revealed to have a dual-lobed structure.

The remainder of the paper is structured as follows. In the next section we provide a review of relevant literature. We then provide data and methods. The results section discusses the impact of choice of geographical unit, an analysis of less populated and rural regions, and the role of population size on a region's position in skills-space. Finally, we draw conclusions and provide tentative policy implications.

Literature Review

Complex systems thinking and complexity economics have continued to gain adoption in a variety of communities. In contrast with the view from neoclassical economics that the economy is a "perfectly humming machine", complexity economics takes the perspective that the economy is more akin to an adaptive ecology of ever-changing networks of interactions (Arthur, 2021).

Analysis of regional economies from a complexity perspective has flourished recently with a burst of applied work using matrices of interactions recast as networks. Relatedness, which measures the connections between various activities, and complexity metrics, which examine prevalence of various economics activities through their interactions, are two veins of applied network research currently being unpacked to examine their causes and consequences (César A. Hidalgo, 2021).

Relatedness measures work to derive information on how various activities relate to one another by way of inferring relations in an agnostic manner (César A. Hidalgo, et al., 2018). For example, one may relate exports of a country by how frequently they are both exported from a country. In this case, the network would be defined by connections between exports. The exports would define nodes in the network while the relationships derived from co-exporting patterns would characterize the edges or arcs. In any such analysis, there are two basic components, the activities under examination and relationship between the activities. With just two underlying components of networks, researchers have been able to map numerous "spaces" using network methods with a complexity perspective.

The product space was among the first to gain widespread notoriety. The seminal paper by Hidalgo et al (2007) builds a product space on the idea that products are related if they are exported frequently together. This is rests on the assumption that the underlying capabilities needed to produce both products are present, and thus the relationship is built on similar capabilities.

Both technology and research spaces have been examined using relatedness approaches. Kogler et al (2013) map the technology space by analyzing how co-classification of patents locate in U.S. cities. Also using patent data, Boschma et al (2015) find that a technology is more likely to enter cities that already have related technologies, with the relatedness of two technologies based on the probability a city patents in one given they patent in another. Rigby (2015) also examines entry of U.S. cities into patenting classes but defines the relationships between patent classes based on patent citations. Away from patenting, Guevara et al (2016) examine the research space by building relationships between fields based on the probability that authors publish in both fields.

The relatedness approach has also been applied to analyze occupation space. Muneeperakul et al (2013) relate occupations to one another if they co-occur as specialization in U.S. Metropolitan Statistical Areas (MSAs) more frequently than would be anticipated at random. These occupation networks were then subsequently used to examine how urban areas might be able to transform their economies into more creative economies (Shutters, Muneepeerakul, & Lobo, 2016) or green economies (Shutters, Muneepeerakul, & Lobo, 2015a).

There have also been a number of papers analyzing the industry space. To map Swedish industry space, Neffke et al (2011) relate industries by the co-occurrence of products being manufactures at the plant level. More recently Waters and Shutters (2020) map industry space by relating industries to one another by how frequently they co-occur as specialized industries based on employment location quotients in U.S. MSAs.

Finally, there has also been work mapping skills space. Neffke et al (2013) examine skill-relatedness of industries by examining inter-industry labor flows. Abdullakareem et al (2018) examine the co-occurrence of skills within occupations to map the skills space. Finally, Shutters and Waters (2020) map the skills space by examining the co-occurrence of specialized skills at the region level by examining U.S. metro areas.

In addition to building the “spaces” of economic activity, there has been a variety of analyses that examine how regional economies are situated in these networks. For example, papers mapping the technology space have determined how likely it is that a city will begin patenting in a field given their current patenting activity. There has also been work on where cities locate within occupation space to provide a sense of how difficult it may be for the region’s economy to transition.

Such locations within economic spaces have also been aggregated to describe the overall inter-connectedness of regional economies. Shutters et al (2015b) aggregate the inter-dependence of occupations based on occupational co-occurrence into a measure called “tightness”, which is intended to capture the share of the region’s economy that is inter-linked. This tightness measure has recently been extended into regional skills tightness (Shutters & Waters, 2020) as well as regional industry tightness (Waters & Shutters, 2020).

While the relatedness concept has been applied widely, there has yet to be attention given to rural regions. While some studies control for regional population regions are typically MSAs and any attention on regional size is peripheral. It is the focus of this paper to analyze how population impacts some measures of relatedness, specifically skills space and industry space. In this sense, this work is well placed in the body of developing research that is currently unpacking the causes and consequences of relatedness (César A. Hidalgo, 2021).

Data and Methods

The following analysis examines more carefully two recent measures, skills tightness, and industry tightness. Before discussing the tightness measure analyzed in the next section, we provide a description of the three primary datasets used. We also provide a brief description of the geographic definitions and the sources of the regional performance indicators used to assess interdependence and tightness.

Data

There are three datasets from two sources. For skills tightness we match occupational employment data to a survey of occupational requirements needed to perform occupations. For industry tightness, we use industry employment data. For both sets of employment data, we use county level data estimated by the Indiana Business Research Center (IBRC). These datasets estimate data suppressed by national statistical agencies (Zheng, 2020). Without occupation and industry employment at county level, we would be unable to examine less populated and rural regions. Such data limitations are likely to add to the list of reasons why rural regions have received less attention than more populous regions.

For the skills tightness measure, we use the occupation employment estimates at the county level by IBRC. Previous analysis of the skills tightness metric used Occupational Employment Statistics (OES) data from the Bureau of Labor Statistics (BLS). OES data are published annually and include employment estimates using Standard Occupation Classifications (SOCs) for U.S. MSAs. A limitation with the OES data is that they only cover Metropolitan Areas and do not cover Micropolitan Areas. Furthermore, they use an alternate regional definition in the New England States of the U.S., known as New England City and Town Areas (NECTAs). The focus of previous analysis could have only occurred for more populous regions. The occupation employment data estimated by IBRC allows for the examination of less populous Micropolitan Areas as well as rural regions. IBRC aggregates several counties in a similar manner to the Bureau of Economic Analysis (BEA), which collapses some counties in Virginia. All data crosswalk county-level accordingly. Finally, the IBRC does not estimate Essex County, VA.

The second data set required for the skills tightness analysis is the Occupational Information Network (O*Net) data from the BLS. This dataset characterizes occupations by several elements, or attributes that are required to perform any given job. Elements included in the O*Net dataset include “Oral Comprehension”, “Design”, “Repairing”, and “Equipment Maintenance”. The BLS performs a survey in order to measure the level and importance of each element associated for each occupation. For this study, we use the level associated with each element. We refer to these elements as skills. Thus, after matching O*Net data to employment data, we are able to build a skills space.

Two notes about conformity between IBRC OES estimate and O*Net data are important. First, because O*Net data does not provide information on legislatures, these SOC codes are dropped from the IU OES estimates. Second, IU estimates several 6-digit SOCs that do not correspond to SOCs in the O*Net dataset. For these “other” categories, we take an average of all elements for 6-digit SOCs in the 4-digit SOC category that they are categorized in with a few exceptions where a subset of the 4-digit SOC was used. Employment in these categories is relatively small and thus is likely to have minimal influence on the analysis.

For the industry tightness we use also use estimated data from the IBRC. These data are employment by place of work at the county level. From the IBRC dataset, we use data classified using the 4-digit North American Classification System (NAICS). As with occupation data, IBRC aggregates several counties in a similar manner to the BEA, which collapses additional counties in Virginia. For all analysis we drop the “balance of industry” estimates from the IBRC data.

Spatial units of analysis

To determine effects of choice of geographical unit, we use three different geographic definitions. First, we use the counties as delineated by IBRC.

Second, we use the Core-Based Statistical Area (CBSA) definitions, which aggregate one or more counties based on a minimum number of people in the core county and commuting from the surrounding counties. CBSAs include both Metropolitan statistical areas, which are 50,000 people or more, as well as Micropolitan statistical Areas which have 10,000 to 50,000 people in the urban center. CBSAs exclude 1,302 counties in the U.S. accounting for 5.9% of U.S. population in 2018. We use the September 2018 CBSA definitions from the Office of Management and Budget (OMB). While we create a "non-CBSA" category to calculate the metric, we drop the non-CBSA region following equation 2.2 for analysis as it is an extreme, non-conforming, outlier.

Third, we use the Labor Market Area provided by Fowler & Jensen (2020). These definitions were originally produced by the U.S. Department of Agriculture's Economic Research Service (ERS) and define labor markets that are inclusive of all U.S. counties. We use the updated "OUT10" delineation that repaired discrepancies in the original definitions published by ERS. Counties, while relatively stable geographies, are in reality arbitrary with respect to economic activity. This is problematic given that a theoretical basis of the tightness metric used here was to measure self-contained labor markets, which rarely correspond to county boundaries. The labor market areas are used in order to cover all counties in the U.S. in a form more comparable with the theoretical foundation of tightness.

Methodology

The industry and skills tightness metrics defined here follow nearly identical formulations of Shutters et al (2015b). As an overview, we calculate the relatedness of the economic activity, either skills or industry, and then aggregate relatedness of the economic activity to measure the "tightness" at the three geographic levels.

Skills tightness requires an additional step not needed for industry tightness. Following the notation of Shutters and Waters (2020), we aggregate skills by weighting O*Net data by employment levels for each geography. Formally,

$$s_{i,g} = \sum_{o} l_{i,o} e_{o,g} \tag{1}$$

where l is the level of skill i required for occupation o. This skill level is used to weight employment, e in the geography of interest, g. The total level of skill is thus the sum of skills weighted employment in the geography under consideration.

We then calculate the commonly used location quotient, for both skills and industry. For skills, we calculate the relative abundance of skill i in geography g.

$$LQ_{i,g} = \frac{\left(s_{i,g} / \sum_i s_{i,g} \right)}{\left(\sum_g s_{i,g} / \sum_g \sum_i s_{i,g} \right)} \,. \tag{2.1}$$

For industries, we use the identical formula with altered notion. We calculate the relative abundance of employment in industry k, in geography g.

$$LQ_{k,g} = \frac{\left(e_{k,g}/\sum_k e_{k,g}\right)}{\left(\sum_g e_{k,g}/\sum_g \sum_k e_{k,g}\right)}. \tag{2.2}$$

For both skills and industry, we convert the matrices of geography by activity into a present absence matrix. The binary matrix is 1 where the economic activity (skill or industry) is specialized in the geography and 0 otherwise. We use an LQ threshold of 1 as the cutoff point for specialization. When calculating CBSA tightness, non-CBSA regions are dropped at this point.

With the present-absence matrix, we then calculate interdependence values using the co-occurrence formula. This is a ratio of the probability that two economic activities *a* and *j* (either skills or industries) are both specialized in a geography over the probability that we would expect to see the activities co-occur at random. We subtract 1 to balance the measure around zero.

$$x_{a,j} = \frac{P\left[LQ_{a,g} > 1, LQ_{j,g} > 1\right]}{P\left[LQ_{a,g'} > 1\right]P\left[LQ_{j,g''} > 1\right]} - 1, \tag{3}$$

If the two activities co-occur more frequently that we would anticipate at random, the interdependence metric is greater than zero. If they co-occur less frequently than we would anticipate at random, the measure is less than zero.

Next, we calculate the aggregate measure of economic tightness, as defined by Shutters et al (2015). To do this, we begin by weighting regional economic activity *a* and *j* by their interdependence values, x. This is Formally:

$$L_{a,j,g} = \frac{\left(s_{a,g} + s_{a,g}\right)x_{a,j}}{2\sum_a s_{a,g}} \tag{4}$$

Finally, we average across all activity pairs to generate the tightness value, *T*.

$$T_g = \frac{2}{p_g\left(p_g - 1\right)}\sum_{a<j}^{p_g} L_{a,j,g} \tag{5}$$

where p_g is the total number of economic activities (either skills or industries) in geography g. Given that skills tightness and industry tightness are calculated using differing techniques, and skills are an arbitrary level based on surveys, we normalize tightness as a z-score for comparability.

<u>*Measures of Economic Performance*</u>

For analysis, we compare the tightness metric derived at three different levels with several measures of economic performance. Performance measure data are from the BEA. Performance measure used are gross domestic product (GDP), earnings by place of work, and employment change. For all measure we use the county level BEA data. BEA counties are aggregated in the case of CBSAs and LMAs .

Results

We examine four aspects of tightness relevant to rural regions. First, we examine the impacts of changing the geographic unit on interdependence of skill and industry pairs. Second, we examine correlations between the two measures of economic tightness and regional economic performance. Third, we explicitly examine the influence of rurality on skills and industry tightness by population decile. Finally, we extend the rural analysis by analyzing relationships between population and position within the skills space before examining possible drivers for rural regions.

Geographic Unit and Interdependence

To examine how the choice of geographic unit impacts interdependencies between skill-pairs and industry-pairs correlate with interdependence values calculated using different geographic units. Scatter plots are provided in the supplemental online material (SOM Fig. 1).

Overall, the geographic unit used to calculate skill-pair and industry-pair interdependence values appears to matter little. Pearson correlation coefficients between interdependence of skills between each of the geographies is 0.983 or higher, all significant at the 1 percent level. Pearson correlation coefficients of industry-pairs from the different geographies range from 0.83 to 0.88, all significant at the 1 percent level.

The strong positive correlations for interdependence values calculated using different geographies implies that the values generated are not an artifact of the geographic unit.

Regional Economic Tightness and Regional Economic Performance

To begin, we examine how skills and industrial tightness calculated using three different areal units correlate with population as well as indicators of regional economic performance (Table 1).

While skills tightness is positively, and significantly associated population is with skills tightness at all geographic levels, the strength of the correlation drops from 0.438 at the CBSA level to 0.247 at the LMA level to just 0.084 at the County level. Industry tightness, in contrast, is more highly correlated with population at the CBSA and County level than the LMA level.

In terms of economic productivity, both skills tightness and industry tightness are positively and significantly correlated with both log of GDP per capita and log of workplace earnings, with the larger geographic units of LMA and CBSA generally having stronger correlations. Regarding year-over-year change in log of GDP and employment, the evidence is inconclusive for skills tightness. The evidence is mixed for industry tightness correlations between log GDP per capita change and employment change from 2018 to 2019.

Overall, the correlations between both tightness measures and population vary notably when the geographic unit is altered while correlations between tightness and economic output are relatively stable. The fact that correlations between tightness and log GDP per capita are more stable than tightness and population between geographic levels suggests that the correlation between tightness and economic output is robust.

Table 1. Correlation Coefficients

	Skills Tightness Z-Score			Industry Tightness Z-Score		
	County	LMA	CBSA	County	LMA	CBSA
Log Population	0.084 (0.000)	0.247 (0.000)	0.438 (0.000)	0.208 (0.000)	0.073 (0.068)	0.228 (0.000)
Log GDP Per Capita	0.135 (0.000)	0.192 (0.000)	0.178 (0.000)	0.262 (0.000)	0.193 (0.000)	0.446 (0.000)
Log GDP Percent Change (`18-`19)	-0.084 (0.000)	0.050 (0.210)	-0.091 (0.006)	-0.213 (0.000)	0.019 (0.635)	-0.165 (0.000)
Log Place of Work Earnings Per Capita	0.162 (0.000)	0.298 (0.000)	0.322 (0.000)	0.377 (0.000)	0.349 (0.002)	0.481 (0.000)
Employment Percent Change (`18-`19)	0.029 (0.110)	0.061 (0.126)	0.098 (0.003)	0.076 (0.000)	0.309 (0.000)	0.222 (0.000)

p-value in parentheses

To examine the sole impact of skills tightness and industry tightness on output, models 1 through 12 regress skills and industry tightness onto log of GDP per capita, controlling for population in models 2,4,6,8,10, and 12 (Table 2). Overall, controlling for population has minimal impact on the point estimates for industry and skills tightness. These results provide evidence that both skills and industry tightness measure an aspect of regional economic structure separate from population. Although the estimates tightness are relatively stable, the models have relatively low explanatory power. Consistent with the correlations, the highest R^2s are found when examining industry tightness and log GDP per capita resulting from CBSAs.

Table 2. Regression Table – Dependent is Log GDP Per Capita

	County				LMA				CBSA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Industry Tightness (Z)	0.154 (0.000)	0.164 (0.000)			0.069 (0.000)	0.069 (0.000)			0.158 (0.000)	0.150 (0.000)		
Skills Tightness (Z)			0.079 (0.000)	0.081 (0.000)			0.069 (0.000)	0.072 (0.000)			0.063 (0.000)	0.042 (0.001)
Log Population		-0.034 (0.000)		-0.016 (0.024)		0.002 (0.783)		-0.005 (0.558)		0.026 (0.003)		0.038 (0.000)
R^2	0.069	0.076	0.018	0.020	0.037	0.037	0.037	0.037	0.199	0.207	0.032	0.047

p-values in parentheses

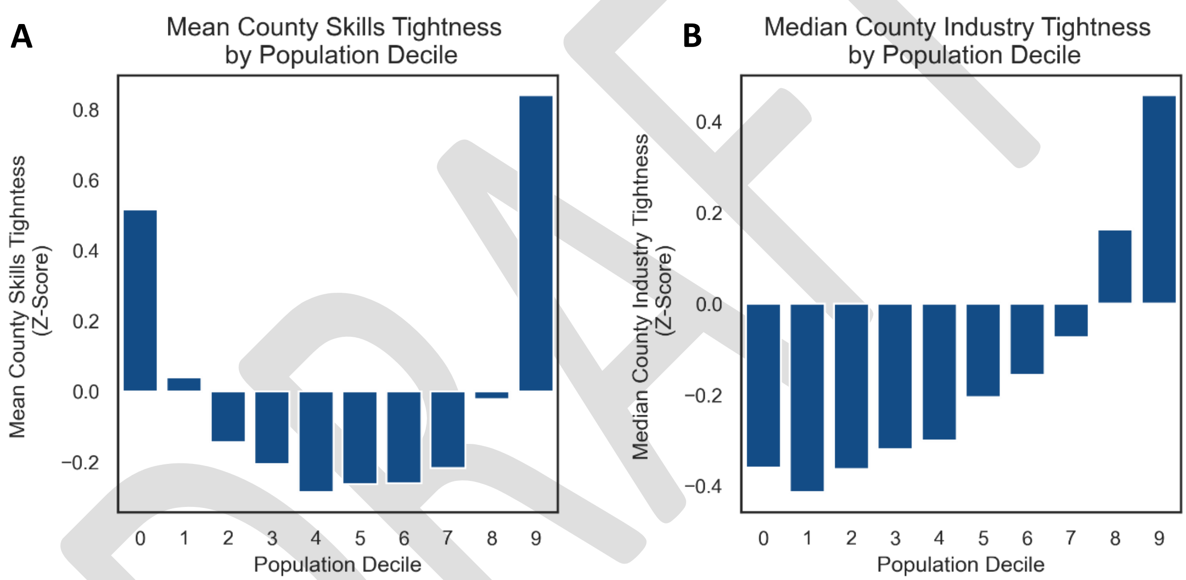
Rurality and Tightness

To analyze associations between rurality and tightness more explicitly, we examine counties by population deciles. In contrast to the conflicting results from the correlations and regressions, examining skills tightness and industry tightness using the population deciles yields clear results (Figure 1). Counties in the bottom and top deciles by population have the highest average within decile skills tightness. Furthermore, there is a relatively smooth transition from the lowest decile to the highest decile, with those counties in the fifth decile having the lowest average within skills tightness.

Examining industry tightness by population decile also yields clear results, although the correlation is more linear, as opposed to non-linear revealed using skills. Median is used for industry tightness due to extreme outliers in the most rural categories. Counties in the second lowest decile have the lowest median tightness score. Counties in the top decile have the highest median industry tightness.

The same general pattern holds when examining LMAs and CBSAs (SOM Fig. 2). The largest LMAs and CBSAs have the highest industry and skills scores. While the smallest LMAs also have a higher mean skills tightness than mid-sized LMAs, this result does appear when using CBSAs. The disappearance of higher tightness in the smallest decile is likely due to the smallest CBSAs being notably more populous than the smallest LMAs. The average population of the smallest CBSA is 21,959 while the average population of the smallest LMA is just 8,299.

Figure 1. Counties by Population Decile and Tightness

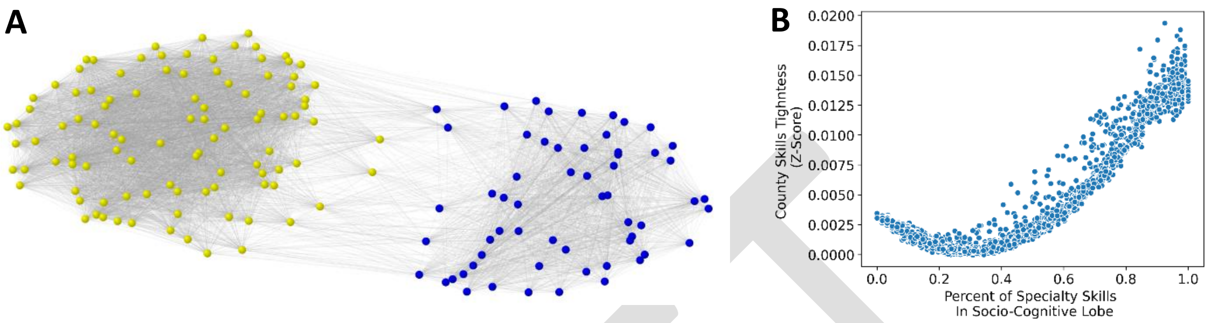


<u>*Rurality and Skills Space Location*</u>

Further focusing on rurality, we analyze possible relationships between county population deciles and location in skills space. As has been previously found, when conceived as a network, the skills space takes a dual-lobed structure (Figure 2A). The two lobes of the network have been termed “Socio-Cognitive” (Yellow) and “Sensory-Physical” (Blue). This structure was not found using industry inter-dependence (Waters & Shutters, 2020). We find the same dual lobed structure here using all three geographies (SOM Fig. 3).

To examine the impact of population on the “location” of geographies in these networks, we compare the percent of a county’s specialty skills in the socio-cognitive lobe to skills tightness (Figure 2B). Each observation characterizes the counties skill tightness and the share of the county’s specialty skills (from equation 2.1). The relationship between county-level skills tightness and the percent of their specialty skills are highly correlated, albeit non-linear. Overall, the higher share of specialty skills in the socio-cognitive lobe, the higher the county's skills tightness.

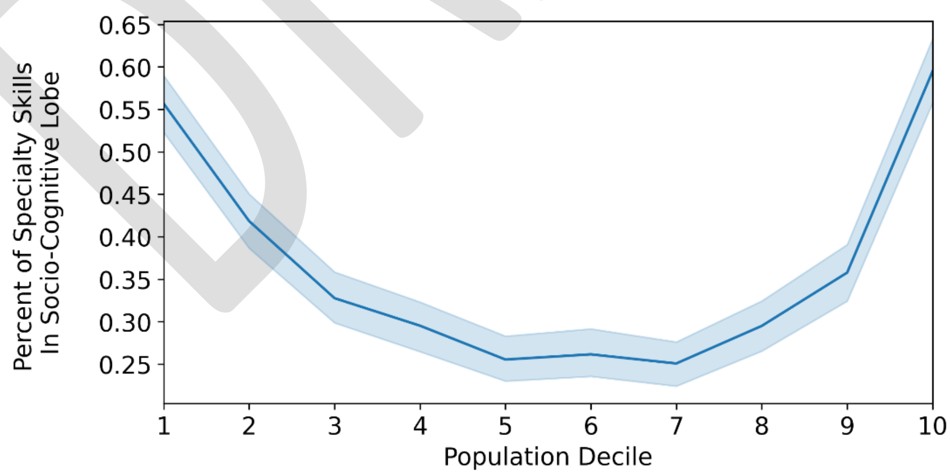
Figure 2. County Based Skills Network



The fact that higher tightness is associated with a greater percent of a county's specializations in the socio-cognitive lobe coupled with the high average tightness of the most rural counties suggests that the most rural regions have a relatively large share of their specialty skills in the socio-cognitive lobe.

To examine the possibility, the average share of skills by population decile is presented in figure 3. The average county in the lowest decile by population has a *higher* average portion of their specialty skills in the socio-cognitive lobe than the average county in the highest decile by population. This finding suggests that the most rural regions may have an advantage in increasing regional productivity through skills tightness relatively to moderately populous counties and perhaps even metro areas.

Figure 3. Share of a County's Specialty Skills in Socio-Cognitive Lobe by Population Decile



Given general differences between the structure of economies in rural and more populous areas, it may be the case that less populous regions specialize in different socio-cognitive skills than more populous counties. Examining the location of the skills in the socio-cognitive lobe revealed little differences between counties in the top and bottom deciles by population. However, examining the most numerous specialty skills in the least and most populous places provides some indication of a difference. The least populous counties appear to have skills geared towards education and training while the most populous appear to have more skills required in application of occupations (SOM Tables 1 & 2). The most numerous specialty skills in counties in the first population decile include skills such as "Education and Training", "Coaching and Developing Others", and "Guiding and Developing Others". Additional skills in the top ten include broad academic topics such as "Geography", "History and Archeology", "Fine Arts", and "Philosophy and Theology". In contrast, the most numerous specialty skills in counties in the top population decile include applied skills such as "Customer and Personal Service", "Service Orientation", "Speech Recognition", "Performing for or Working Directly with the Public", "Negotiation", "Coordination", and "Persuasion". These skills suggest that rural and more populous places may be performing notably different tasks within the economy, despite both being located in the socio-cognitive lobe.

Finally, to decipher which economic tasks differ between rural and urban regions, we examine the contribution to socio-economic skills in the bottom and top decile by occupation (SOM Table 3). Specifically, we examine the matrix produced from equation 1 where skill *i* is in the socio-cognitive lobe and geography, g, are the counties in either the bottom or top decile by county. We simply examine the difference of occupation contribution to socio-cognitive skills between the least and most populous counties.

Examining rural counties first, we see that the four of the largest differences in contribution to skills are from teachers. That is, teachers contribute disproportionately more socio-cognitive skills in the least populous counties. This is consistent with the findings by skills which suggested higher training activity in rural regions. We also see notably higher contributions by farming occupations such are truck drivers, farmworkers, and meat and poultry trimmers.

The occupations contributing disproportionately the most to socio-cognitive skills in the most in populous regions tend to be service sector jobs such as retail salesperson, personal care aides, home health aides and security guards. What is important to point out is that most of the occupations that contribute disproportionately to the socio-cognitive lobe in either the most or least populous counties appear to be support occupations. That is, they are not basic economic activity, but rather support the primary economic activity. Notable exceptions, such as the farming operations mentioned, are however primarily rural activities. These occupations, while perhaps not commonly associated with socio-cognitive activity contribute to the both the socio-cognitive activity in rural regions, helping to increase both skills tightness as well as regional economic output.

Conclusion

As analysis of regional economies as complex systems has continued to grow, less populated regions have been overlooked. This is likely due to data limitations as well as findings in the scaling literature. Despite this focus, less populated and rural regions are nonetheless complex systems, with much to be learned by analyzing them in this manner. This paper examines rural regions and complex adaptive systems by calculating a metric of economic tightness at three geographic levels as specifically examining the impact of population.

In the paper we uncovered important results with respect to rural regions and economic tightness. First, we found the geographic unit used to build interdependence values between skills or industries doesn't impact interdependence values. Second, tightness impacts economic output and is not influenced heavily by regional population. Most importantly, we find that the most rural and most urban regions have the highest share of their specialty skills in the socio-cognitive lobe of the skills-space, coinciding with higher skills-tightness. Given that skills-tightness is positively correlated with economic output, the most rural regions may have structural advantages with respect to skills tightness.

Finally, we tentatively find that there may more teaching and training skills in the most rural regions and more applied skills in more populous regions. Apart from what are perhaps non-basic economic activities, we find that farm activity contributes notably to the socio-cognitive lobe of the skills space in the least populous counties. This suggests that such occupations should not be overlooked by regional economic developers who are planning to diversify into socio-cognitive economic activity.

Supplementary Materials: Included with this pdf file.

Author Contributions: Conceptualization, K.W. and S.T.S.; methodology, K.W. and S.T.S.; investigation, K.W. and S.T.S.; writing—original draft preparation, K.W. and S.T.S.; writing—review and editing, K.W. and S.T.S.; visualization, K.W. and S.T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This publication was prepared by Keith Waters at George Mason University and Shade T. Shutters at Arizona State University using Federal funds awarded to the Trustees of Indiana University and as a sub-component under award number ED17HDQ3120040 from the U.S. Economic Development Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the Economic Development Administration or the U.S. Department of Commerce.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking the polarization of workplace skills. *Science Advances, 4*, eaao6030.

Arthur, W. B. (2021). Foundations of complexity economics. *Nature Reviews Physics, 3*, 136-145.

Boschma, R., Balland, P.-A., & Kogler, D. F. (2015). Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change, 24*, 223-250.

Fowler, C. S., & Jensen, L. (2020). Bridging the gap between geographic concept and the data we have: The case of labor markets in the USA. *Environment and Planning A: Economy and Space, 52*, 1395-1414.

Guevara, M. R., Hartmann, D., Aristarán, M., Mendoza, M., & Hidalgo, C. A. (2016). The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations. *Scientometrics, 109*, 1695-1709.

Hidalgo, C. A. (2021). Economic complexity theory and applications. *Nature Reviews Physics*.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., & Zhu, S. (2018). The Principle of Relatedness. In A. J. Morales, C. Gershenson, D. Braha, A. A. Minai & Y. Bar-Yam (Eds.), *Unifying Themes in Complex Systems IX* (pp. 451-457). Cham: Springer International Publishing.

Hidalgo, C. A., Klinger, B., Barabasi, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science, 317*, 482-487.

Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies, 21*, 1374-1391.

Muneepeerakul, R., Lobo, J., Shutters, S. T., Goméz-Liévano, A., & Qubbaj, M. R. (2013). Urban Economies and Occupation Space: Can They Get "There" from "Here"? *Plos One, 8*, e73676.

Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal, 34*, 297-316.

Neffke, F., Henning, M., & Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography, 87*, 237-265.

Rigby, D. L. (2015). Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes. *Regional Studies, 49*, 1922-1937.

Shutters, S. T., Muneepeerakul, R., & Lobo, J. (2015a). How hard is it for urban economies to become "green"? *Environment & Planning B: Planning and Design, 43*, 198-209.

Shutters, S. T., Muneepeerakul, R., & Lobo, J. (2015b). Quantifying urban economic resilience through labour force interdependence. *Palgrave Communications, 1*, 1-7.

Shutters, S. T., Muneepeerakul, R., & Lobo, J. (2016). Constrained pathways to a creative urban economy. *Urban Studies, 53*, 3439-3454.

Shutters, S. T., & Waters, K. (2020). Inferring Networks of Interdependent Labor Skills to Illuminate Urban Economic Structure. *Entropy, 22*, 1078.

Slaper, T. (2019). Rural Economic Development, Wicked Problems and Complex Adaptive Systems. *Indiana Business Review, 94*, https://www.ibrc.indiana.edu/ibr/2019/fall/article2012.html.

Waters, K., & Shutters, S. T. (2020). Industrial structure and a tradeoff between productivity and resilience. *SSRN Electronic Journal, https://doi.org/10.2139/ssrn.3702545*.

West, G. (2017). *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. New York: Penguin Press.

Zheng, P. (2020). Filling in the holes: Generating point estimates for QCEW suppressed data. *InCotext, 21*, http://www.incontext.indiana.edu/2020/jan-feb/article2022.asp.

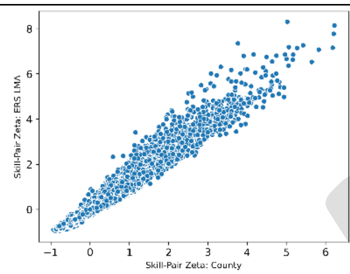
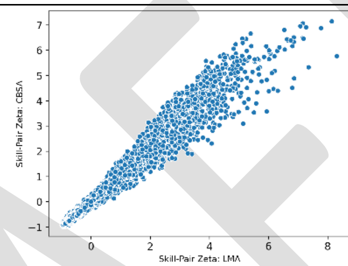
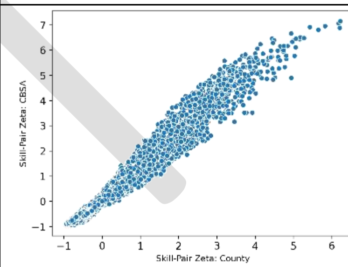
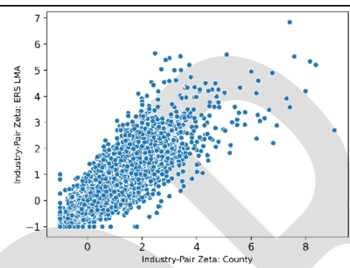
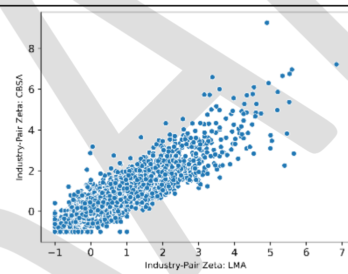
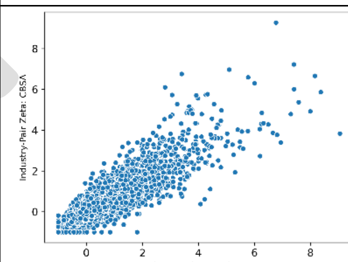
Supplemental Online Material

Keith Waters and Shade T. Shutters

Interdependence Values of Skill-Pairs and Industry-Pairs Correlations

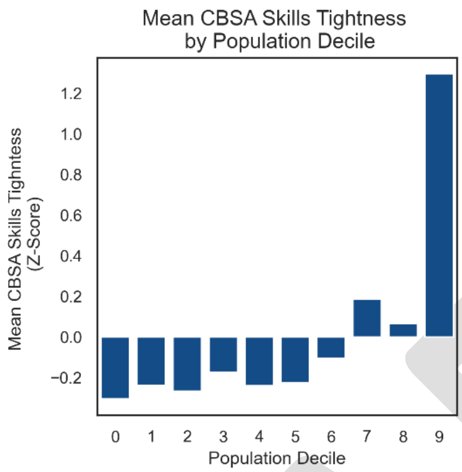
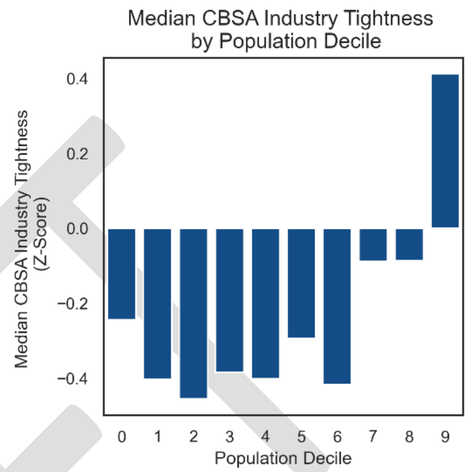
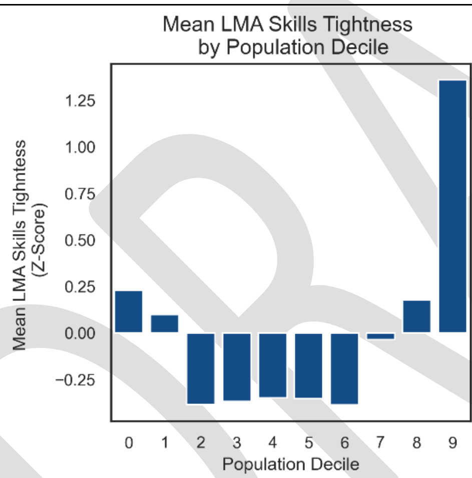
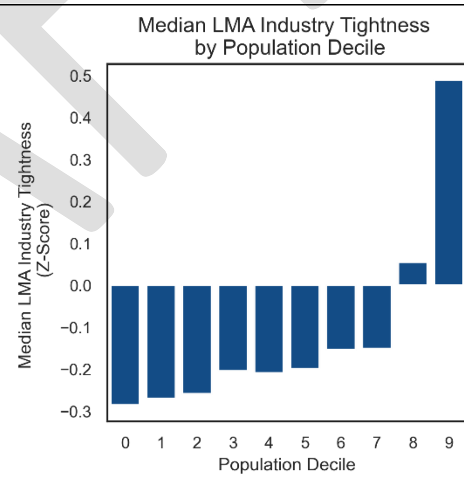
In each scatter plot, each data point is a skill-pair or industry pair, and each axis is the skill/industry-pair interdependence value calculated using the noted geography.

SOM Figure 1. Interdependence Values Calculated Using Different Geographies

	County and LMA	ERS-LMA and CBSA	County and CBSA
Skill-Pairs			
Industry-Pairs			

LMA and CBSA Skills Tightness

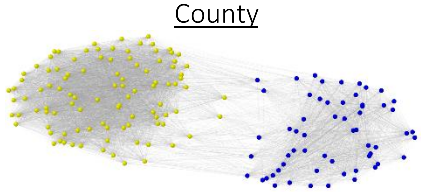
SOM Figure 2. LMAs and CBSAs by Population Decile and Tightness

<table>
<tr><th></th><th>Mean Skills Tightness (Z-Score)</th><th>Median Industry Tightness (Z-Score)</th></tr>
<tr><td>CBSA</td><td>
</td><td>
</td></tr>
<tr><td>LMA</td><td>
</td><td>
</td></tr>
</table>

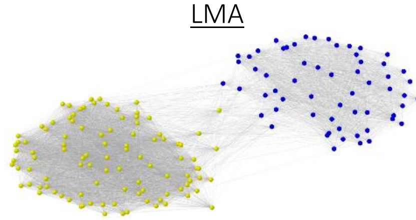
Skills Networks

SOM Figure 3. Skills Network Dual Lobed Structure at County, LMA, and CBSA

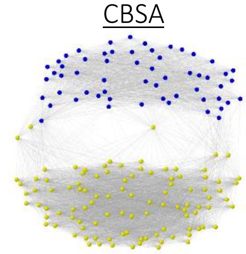
County



LMA



CBSA



Socio-Cognitive Skills of Rural Counties vs Populous Counties

SOM Table 1. Socio-Cognitive Specialty Skills of Counties in the Bottom Decile

Rank	**Element (Skill)**	**Element (Skill) Name**	**Population Decile**	**Counties with Specialty Skill**
1	2.C.4.g	Geography	1	260
2	2.C.7.d	History and Archeology	1	247
3	2.C.6	Education and Training	1	234
4	4.A.4.b.3	Training and Teaching Others	1	230
5	4.A.2.b.5	Scheduling Work and Activities	1	223
6	4.A.4.b.5	Coaching and Developing Others	1	211
7	4.A.4.b.4	Guiding, Directing, and Motivating Subordinates	1	205
8	2.C.7.c	Fine Arts	1	203
9	2.C.7.e	Philosophy and Theology	1	201
10	4.A.4.a.3	Communicating with Persons Outside Organization	1	200

SOM Table 2. Socio-Cognitive Skills of Counties in the Top Decile

Rank	**Element (Skill)**	**Element (Skill) Name**	**Population Decile**	**Counties with Specialty Skill**
1	2.B.1.f	Service Orientation	10	200
2	2.C.1.e	Customer and Personal Service	10	200
3	1.A.4.b.4	Speech Recognition	10	193
4	4.A.4.a.8	Performing for or Working Directly with the Public	10	192
5	4.A.4.a.7	Resolving Conflicts and Negotiating with Others	10	188
6	2.B.1.d	Negotiation	10	186
7	2.B.1.b	Coordination	10	185
8	4.A.4.a.4	Establishing and Maintaining Interpersonal Relationships	10	183
9	2.B.1.c	Persuasion	10	181
10	1.A.1.d.1	Memorization	10	176

Difference in Contribution to Socio-Cognitive Skills by Occupation Between Rural and Urban Counties

SOM Table 3. Occupational Contribution of Socio-Cognitive Skills for Rural and Populous Counties

	SOC Code	**SOC**	**Decile 1 Percent of LVC 1 Skills**	**Decile 10 Percent of LVC 1 Skills**	**Difference**
More Common in Populous Counties	41-2031	Retail Salespersons	1.8%	3.0%	-1.2%
	35-3021	Combined Food Preparation and Serving Workers, Including Fast Food	1.6%	2.2%	-0.6%
	39-9021	Personal Care Aides	0.8%	1.4%	-0.5%
	43-4051	Customer Service Representatives	1.9%	2.4%	-0.5%
	35-3031	Waiters and Waitresses	1.2%	1.6%	-0.4%
	31-1011	Home Health Aides	0.3%	0.6%	-0.3%
	31-9092	Medical Assistants	0.3%	0.6%	-0.3%
	33-9032	Security Guards	0.5%	0.8%	-0.3%
	35-2014	Cooks, Restaurant	0.7%	0.9%	-0.3%
	43-4171	Receptionists and Information Clerks	0.5%	0.7%	-0.2%
More Common in Rural Counties	51-3022	Meat, Poultry, and Fish Cutters and Trimmers	0.4%	0.1%	0.3%
	45-2092	Farmworkers and Laborers, Crop, Nursery, and Greenhouse	0.5%	0.2%	0.3%
	29-1141	Registered Nurses	3.4%	3.0%	0.4%
	37-2012	Maids and Housekeeping Cleaners	0.8%	0.4%	0.4%
	25-2022	Middle School Teachers, Except Special and Career/Technical Education	0.9%	0.5%	0.4%
	25-9041	Teacher Assistants	1.2%	0.8%	0.4%
	25-2031	Secondary School Teachers, Except Special and Career/Technical Education	1.6%	0.9%	0.7%
	41-2011	Cashiers	2.9%	2.0%	0.8%
	25-2021	Elementary School Teachers, Except Special Education	2.1%	1.2%	0.9%
	53-3032	Heavy and Tractor-Trailer Truck Drivers	2.1%	1.1%	1.0%