Waters and Shutters *Urban Transformations* (2022) 4:15
https://doi.org/10.1186/s42854-022-00044-6

Urban Transformations

RESEARCH **Open Access**

Regional economic tightness from rural to urban regions

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

*Correspondence: keith.waters@asu.edu

[1] School of Complex Adaptive Systems, Arizona State University, Washington, DC, USA
[2] School of Complex Adaptive Systems, Arizona State University, Tempe, AZ, USA

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 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. Smaller urban areas are typically over-looked and rural regions are almost entirely absent from the dialog. This paper seeks to fill this gap by examining smaller urban areas and rural regions from a complexity economics perspective. Analyzing cross-sectional data provides initial insights into the transformation of regional economic connectedness from rural to urban regions. Using a previously developed metric of economic connectivity based the on co-occurrence of economic activities, called tightness, we examine the skills space and industry space of metropolitan, micropolitan, and rural regions in the United States. We find that the least and the most urbanized regions have the highest tightness, and that this is partially due to the share of specialty skills in a "socio-cognitive" lobe of skills space. However, we also find that the composition of skills in the least and most urbanized regions differs markedly. Findings suggest that planners seeking to increase the share of socio-cognitive skills in the local workforce may be constrained by population size, and that regions of moderate population size may be required to first grow industries that require less cognitive skills.

Science Highlights

• Regional economic tightness and regionaleconomic output are positively correlated, even when controlling for regionalpopulation.

• Skills tightness is greatest in the leastpopulous and most populous regions while industry tightness is greatest in themost populous regions.

• Higher skills tightness is driven partially bythe share of socio-cognitive skills in the regional workforce.

• The most rural and most urban counties havethe highest share of specialty skills in the socio-cognitive lobe of skillsspace.

Policy and PracticeRecommendations

• Both skills and industrial tightness should befostered to increase regional per capita output.

• Growing jobs that utilize socio-cognitiveskills may increase skills tightness and thus regional productivity.

• Moderately-urbanized areas typically have a lowershare of workers with socio-cognitive skills, and may experience more difficultygrowing knowledge-intensive industries.

Keywords: Complexity, Tightness, Skills, Productivity

Introduction

Work analyzing regional economies from a complex systems perspective 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 urban regions is likely due to data limitations in rural regions as well as the super-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 urban and rural regions are worthy of attention given structural disadvantages resulting from size documented in the scaling literature.

In this paper, we analyze regional economies as complex 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 regional economics, focusing on how regional economic tightness varies from rural to urban regions. Using cross-sectional data, the aim is to provide initial evidence on structural changes that regions face as they transform from rural to urban regions. 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, we systematically assess how choice of areal units affects co-occurrence and aggregate economic tightness, two measures used to examine regional economies as complex systems. Second, we examine the impact of regional economic tightness on regional measures of regional economic performance before controlling for population. Specifically focusing on population allows for initial evidence regarding tightness and rural to urban transformation. Finally, we analyze 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. In the results section we discuss 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 provide a discussion of the implications for planners before drawing conclusions.

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 applied work recasting interactions 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 (Hidalgo 2021).

Relatedness measures attempt to determine how various activities relate to one another by inferring relations in an agnostic manner (Hidalgo et al. 2018). For example, one may relate two products by how frequently they are both exported from various countries. In this case, the network is defined by connections between products. The exported products would define nodes in the network while the relationships derived from co-exporting patterns would characterize the edges in the network. In any such network analysis, there are two basic components, the activities under examination and relationship between the activities. Using these networks, researchers have mapped numerous "spaces".

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 two products are related if they are frequently exported from the same country. Two products that are both frequently exported from the same countries are inferred to require similar underlying capabilities.

Technology and research spaces have since 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. Muneepeerakul et al. (2013) relate occupations to one another if they co-occur as specialization in U.S. Metropolitan Statistical Areas (MSA) 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 et al. 2016) or green economies (Shutters et al. 2015a).

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

Researchers have also used this methodology to map skills space. Neffke and Henning (2013) map the skill-relatedness of industries by examining inter-industry labor flows, while Alabdulkareem et al. (2018) examine the co-occurrence of skills within occupations to map the skills space. Finally, Shutters and Waters (2020a, b) map the skills space using the co-occurrence of specialized skills within U.S. MSAs.

In addition to mapping the spaces of economic activity, there have 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. Careful mapping of such spaces are crucial to advising transitions between portions of the network.

Finally, locations within economic spaces have also been aggregated to describe the overall inter-connectedness of regional economies. (Shutters et al. 2015b) aggregate the interdependence 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. The tightness measure has recently been extended into regional industry tightness (Shutters and Waters 2020a) as well as regional skills tightness (Shutters and Waters 2020b).

While the relatedness concept has been applied widely, there has yet to be attention given to rural regions (Fig. 1). 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 (Hidalgo 2021).

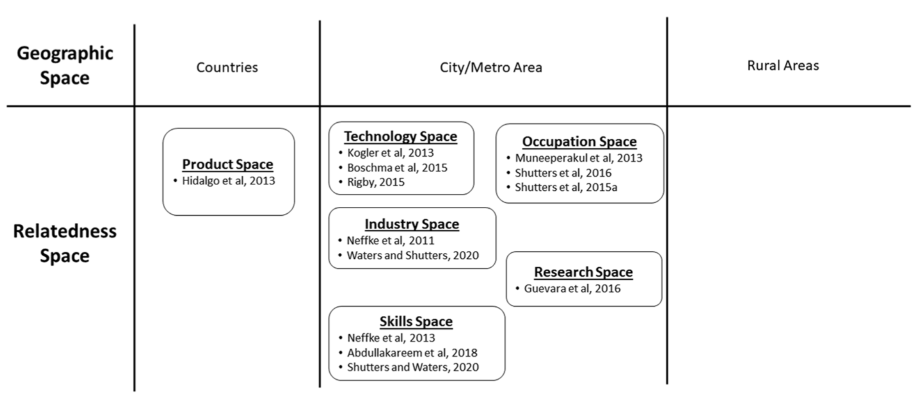

Fig. 1 Conceptual Map of Relatedness Literature. While relatedness measures from complexity theory have measured economic activity at the country level and metropolitan level, rural areas have not received attention in the literature

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 and a brief description of the geographic definitions. After defining the tightness measures, we provide the source of regional performance indicators used to assess tightness.

Data

We use 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 suppressions is perhaps the primary reason why rural regions have not received the attention given to urban regions.

For the skills tightness measure, we use the occupation employment estimates at the county level created 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 (U.S. Bureau of Labor Statistics 2018). A limitation with the OES data is that they only cover Metropolitan Statistical Areas and do not cover Micropolitan Statistical Areas, which are smaller. 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 analyses could have only occurred for more urban regions. The occupation employment data estimated by IBRC allows for the examination of less urban Micropolitan Statistical Areas as well as rural regions. The occupational data from the IBRC is an augmented version of the OES data, in which federally suppressed data is estimated and added. IBRC aggregates several counties in a similar manner to the Bureau of Economic Analysis (BEA), which collapses some counties in Virginia. 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 (National Center for O*NET Development 2020). 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". O*Net conducts 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 skills information on legislatores, 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 IBRC data which estimates industry employment by place of work at the county level using the 4-digit North American Industry Classification System (NAICS) code. The data from the IBRC is an augmented version of the Quarterly Census of Employment and Wages (QCEW) from the US Bureau statistics (U.S. Bureau of Labor Statistics 2022), in which federally suppressed data is estimated and added. As with occupation data, IBRC aggregates several counties in a similar manner to the BEA, which collapses additional counties in Virginia. For all analyses 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 Core-Based Statistical Areas (CBSA), which aggregate one or more counties based on a minimum number of people in the core county and commuting from the surrounding counties (U.S. Office of Management and Budget 2020). 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 U.S. Office of Management and Budget (OMB). While we create a "non-CBSA" category to calculate the metric, we drop the non-CBSA region following Eq. 2.2 for analysis as it is an extreme, non-conforming, outlier.

Third, we use Labor Market Areas (LMA) defined by Fowler and 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. Labor Market Areas are thus used in order to cover all counties in the U.S. and to conform to the theoretical foundations of the tightness calculation.

Methodology

The industry and skills tightness metrics defined here follow nearly identical formulations of (Shutters et al. 2015b). As an overview, we calculate the pairwise interdependence of economic activities, either skills or industries, and aggregate across activities to measure the "tightness" for each of the three geographic levels.

Skills tightness requires an additional step not needed for industry tightness. Following the notation of Shutters and Waters (2020a, b), 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 presence-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 presence-absence matrix, we then calculate interdependence values using the co-occurrence formula. This is the probability that two economic activities a and j (either skills or industries) are both specialized in a geography divided by the probability that activities co-occur randomly. 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 two activities co-occur more frequently than would be expected at random, the interdependence metric is greater than zero. If they co-occur less frequently than expected at random, the measure is less than zero. Thus, an interdependence value at or near zero indicates that the co-occurrence pattern of two activities is essentially random.

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

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

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.

Measures of economic performance

For analysis, we compare the tightness metric derived at three different levels with several measures of economic performance. Performance data are taken from the BEA and include gross domestic product (GDP), earnings by place of work, and employment change. For all measures we use the county level BEA data. BEA county data are aggregated to 2018 CBSAs and the updated ERS LMAs for analysis.

While we only measure economic output, there a number of areas this work could be extended to. For example, it is commonly noted that GDP is not necessarily an indication of improved well-being, and a variety of "beyond GDP" measures that have been analyzed may prove fruitful (Cavalletti and Corsi 2018). For a complete list of such indicators, see (OECD 2013). Similarly, while the present work only examines economic efficiency, it could be extended to examine economic resilience in various difference ways (Martin and Sunley 2015). Finally, this work could be further extended into the area of sustainability, which may at times be at odds with resilience (Elmqvist et al. 2019). For the present analysis, we restrict ourselves to GDP, earnings, and employment.

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 population 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 our analysis, we compare interdependence values calculated using different geographic units. Scatter plots and correlations are provided in the supplemental online material (SOM Fig. 1).

Overall, the choice of geographic unit has little effect on skill-pair and industry-pair interdependence values. Pearson correlation coefficients between interdependence of skills under each pair of the geographies is 0.983 or higher and are all significant at the 1 percent level. Pearson correlation coefficients of industry-pairs under 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

Turning to examine the relationship between tightness and economic performance, the correlation between tightness, both skills and industrial using three different geographic units, and population as well as indicators of regional economic performance are provided in Table 1.

While skills tightness is positively and significantly associated with the log of population 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. The correlation coefficient between industry tightness and population, in contrast, is greater 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.

To examine the impact of skills tightness and industry tightness on output controlling for population, 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 of tightness are relatively stable, the models have low explanatory power. Model 10 has the highest R^2 of 0.207.

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

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

Population and tightness

To analyze associations between population and tightness explicitly, we examine counties by population deciles. In contrast with conflicting results from the correlations and regressions, examining skills tightness and industry tightness by population deciles yields clear results (Fig. 2). Counties in the bottom and top deciles by population have the highest average within decile skills tightness (Fig. 2A). The smallest counties by population have a high average skills tightness. That is, skills of the local workforce are relatively interdependent. Counties in the 4th through 6th deciles have the lowest average skills tightness, indicating that skills in the local workforce are less interdependent than those in the lowest and highest decile. The most populous counties have on average the highest skills tightness. This indicates that the skills in the largest counties are the most interdependent. 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 skills tightness. While the analyzed data are cross-sectional, these data suggest that as counties transform from rural to urban, they go through a transition in which skills tightness decreases.

In practical terms, this implies that the most rural counties have economies with highly interdependent skills. Examining the next largest regions reveals that the sets of skills present in the region are less interdependent on one another. Finally, the most populous counties have the most interdependent skill sets. One plausible explanation for this is that the economic activity in counties in the middle decile have not coalesced around an economic activity. That is, middle decile counties are perhaps still searching for a defining economic activity that binds the regions skills together. While counties in these deciles may have highly skilled labor forces, the skills appear to be disjointed. In the event of an economic shock, disjointed skills may result in an outflow of workers migrating to a region with complementary skills. However, such disjointed activity may also provide economic resilience as such counties don't have their skills entirely focused on an economic activity that experiences a shock.

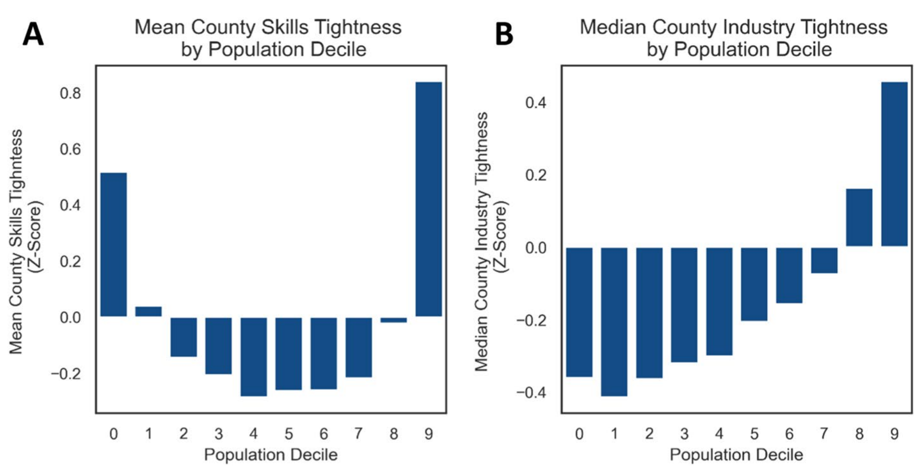

Fig. 2 Counties by Population Decile and Tightness. Skills tightness is greatest for the most rural and the most urban counties, with moderately urban counties (deciles 4, 5, and 6) having the lowest skills tightness (**A**). Industry tightness is greatest for the most urban economics (**B**)

Examining industry tightness by population decile also yields clear results (Fig. 2B). 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. These cross-sectional data suggest that county tightness increases nearly monotonically as population increases.

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 not appear when using CBSAs. The disappearance of higher tightness in the smallest decile is likely due to the smallest CBSAs being notably more urban 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.

Population and skills space location

Further focusing on population, we analyze possible relationships between county population deciles and location in skills space. It has been previously found that when skills are conceived as a network, the skills space takes a dual-lobed structure (Alabdulkareem et al. 2018). That is confirmed here (Fig. 3A). That is, there are two distinct communities of the network that are relatively detached from one another but are internally highly connected. The two lobes of the network have been termed “Socio-Cognitive” (Yellow) and “Sensory-Physical” (Blue). The socio-cognitive lobe of the network are skills that are either social or highly cognitive in nature. For example, three skills that appear in the sensory-physical lobe of the county network below include "Visualization", "Stamina", and "Building and Construction". Three skills that appear in the socio-cognitive lobe include "Economics and Accounting", "Geography", "Analyzing Data or Information". This structure was not found using industry inter-dependence (Shutters and Waters 2020a). We find the same dual lobed structure here using all three geographies (SOM Fig. 3). These findings imply that socio-cognitive skills tend to co-locate with one another in geographic space, whether the geographic unit examined is county, LMA or CBSA. Furthermore, sensory-physical also tend to collocate with one another. If regional planners are looking to transform the skills of their region’s economy, it may be fruitful to diversify within the lobe that their region’s economy is predominately located

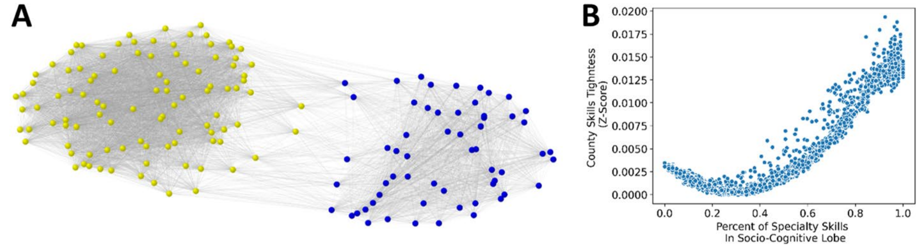

Fig. 3 County Based Skills Network. The skills network, where nodes are equivalent to skills and edges are interdependence scores, *x,* is comprised of two lobes, a socio-cognitive lobe and a sensory-physical lobe (**a**). The percent of a county’s specialty skills in the socio-cognitive lobe is correlated with greater skills tightness (**B**)

in. Transforming from a predominately sensory-physical economy to a socio-cognitive may neither be attainable nor advisable if there are not complementary skills present in the region. The ability to transform regional economies to creative economies (Shutters et al. 2016) and green economies (Shutters et al. 2015a) and been previously explored.

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 (Fig. 3B). Each observation characterizes the county's skill tightness and the share of the county's specialty skills (from Eq. 2.1). The relationship between county-level skills tightness and the percent of the county's specialty skills are highly correlated, albeit non-linearly. Overall, the higher share of specialty skills in the socio-cognitive lobe, the higher the county's skills tightness.

The fact that higher tightness is associated with a greater share 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 this possibility, the average share of skills by population decile is presented in Fig. 4. As Fig. 3 reveals that tightness and the share of socio-cognitive skills is positively correlated, it is unsurprising that Fig. 4 is broadly similar to Fig. 2A, which showed the average tightness by population decile. The most rural and the most urban counties have, on average, a *higher* portion of their specialty skills in the socio-cognitive lobe than moderately urban counties.

While the underlying forces resulting in moderately urban counties to have a lower share of socio-cognitive skills are unclear, the lower share of socio-cognitive skills is likely to be the root cause of lower skills tightness in these counties. Thus, moderately urban counties working to increase the economic tightness of their region in order to become more productive may experience difficulties incentivizing socio-cognitive economic activities. There are at least two plausible reasons for the lower share of socio-cognitive skills in moderately urban counties. First, socio-cognitive skills may benefit from proximity to as many other socio-cognitive activities as possible. Proximity to diversity could incentivize such activities to agglomerate in the most populous counties. Second, sensory-physical activities may have greater space requirements, as in goods

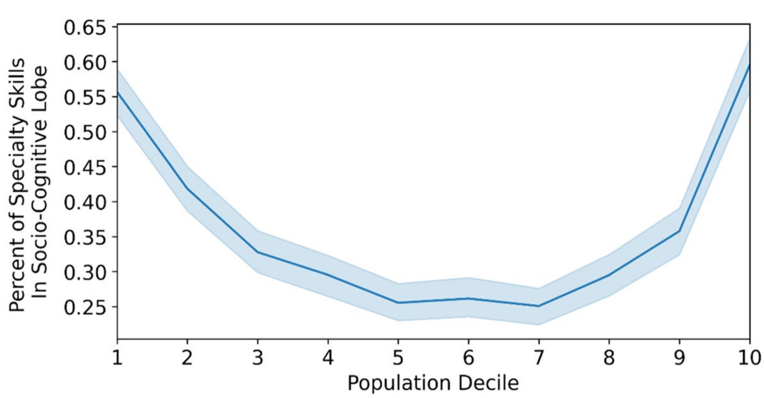

Fig. 4 Share of a County's Specialty Skills in Socio-Cognitive Lobe by Population Decile. Counties in the lowest and highest deciles have the greatest share of their specialty skills in the socio-cognitive lobes. Moderately urban counties, those in deciles 5, 6, and 7, have the lowest share of their specialty skills in the socio-cognitive lobe

manufacturing as well as distribution. Such space requirements may result in sensory-physical activities locating in moderately urban counties where space is at less of a premium, thus pushing socio-cognitive out of such areas. Such dynamics are worth future exploration in temporal analysis.

Finally, given general differences between the structure of economies in rural and more urban areas, it may be the case that less urban regions specialize in different socio-cognitive skills than more urban 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 urban places provides some indication of a difference. The least urban counties appear to have skills geared towards education and training while the most urban appear to have more skills required in application of skills (SOM Tables 1 and 2). The most numerous specialty skills in the least populous counties 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 urban places may be performing notably different tasks within the economy, despite both being located in the socio-cognitive lobe.

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 Eq. 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 urban counties.

Examining rural counties first, 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 urban 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 urban 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 urban 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.

Discussion

We have used a measure of relatedness based on co-occurrence to infer relations both between different skills and between different industries. These relations were then used to measure the inter-connectedness, or tightness, of regional economies. Skills tightness and industry tightness were found to be correlated with output per capita. While industry tightness generally increases with population, skills tightness was highest on average for the most rural and most populous counties. Examining the skills tightness of counties more in-depth revealed a relationship between the share of a county's specialty skills in the socio-cognitive lobe and the county's skills tightness. The average share of county level specialty skills in the socio-economic lobe by population decile revealed that the most rural and most urban countries had the highest share of specialty skills in the socio-cognitive lobe. While these counties at the ends of the rural–urban continuum had high portions of specialty skills in the socio-cognitive lobe, rural area skills tended to be centered on training, such as "Coaching and Developing Others", while urban area skills tended to center on application, such as "Customer and Personal Service".

The association between the share of a county's specialty skills in the socio-cognitive lobe and county level tightness paired with the correlation between tightness and output per capita, may suggest planners should work to increase the share of economic activity in the socio-cognitive lobe. Increasing specialty skills in the socio-cognitive lobe potentially could increase tightness and ultimately GDP per capita. For example, a workforce planner would first examine which occupations have the highest level of socio-cognitive skills. Then, the planner could either work to further incentivize occupations with skills already specialized in the region, or determine skills that are typically co-located, and thus related in some manner in order to diversify the region while building on known strengths. An application of a similar policy tool is discussed by (Waters and Shutters 2022).

As this analysis is cross-sectional data, expected long-run outcomes are difficult to anticipate. Planners typically seek to retrain re-skilled workers locally so that they contribute to regional economic advancement. However, re-skilled workers may find it more beneficial to migrate to other regions to realize higher returns on their newly acquired skills. Such worker migration is plausible as the moderately urban regions have, on average, the lowest tightness and the lowest share of specialty skills in the socio-cognitive lobe. While such a dynamic may work to reinforce existing disparities between regions, regional planners may wish to make investments to retain high skilled workers such as incentivizing industries that require high skilled workers or developing amenities that are demanded of high skilled workers.

Finally, it is important to note several limitations to this work. First, we use a relatively novel methodology to build our industry and skills spaces. As a highly active area of interdisciplinary research, we acknowledge that there is no generally accepted approach to building these types of network construction and thus, there is ample opportunity for further research into the theoretical foundations of our methods. For instance, co-occurrence analysis has long been used in the field of ecology to measure interactions between species (Gotelli 2000; Veech 2013; Griffith et al. 2016). However, there are alternative measures of interdependence, and these should be compared to those developed recently in economic geography. Second, the measure of tightness used in this study is an attempt to define an aggregate, system-level metric of interdependence, integration, or interaction strength. Similar measures exist such as generalized network density (Tokuyama 2007; Liu et al. 2009; Shutters et al. 2018) and economic gravity models (Gómez-Herrera 2013) should be compared to this tightness metric to determine how each correlates with economic performance measures of interest. Finally, this study examined a cross-section of data. While this provides insight into the current structure of the regional system, to elucidate the dynamics more fully, temporal analysis is needed.

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. 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. This provides initial evidence on transformations regional economies go through regarding tightness as they become more urban.

In the paper we uncovered 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 is correlated with economic output. The correlation between tightness and economic output is robust even when controlling for population.

Perhaps 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. Moderately urban areas, those in the 4^{th}, 5^{th} and 6^{th} population deciles have the lowest share of specialty skills in the socio-cognitive lobe, and therefore may be structurally dis-advantaged with respect to skills tightness. While the forces driving moderately urban areas to have lower shares of socio-cognitive specialty skills are not yet clear, possible reasons for this disadvantage include both push and pull factors. Socio-cognitive skills may be pulled to more populous areas where such skills benefit from near a greater amount of economic activity. Socio-cognitive skills may also be pushed out of moderately sized areas if sensory-physical activities have greater space requirements and outbid socio-cognitive activities for space. Thus, planners working to transition moderately urban economies to more socio-cognitive activities may be at a structural disadvantage and may want to take such results into consideration regarding realistic targets.

Finally, while rural areas have a greater share of socio-cognitive specialty skills than moderately sized cities, the specialty skills found in the most rural places tends to be qualitatively different from those found in the most urban areas. While urban areas tend to have specialty socio-cognitive skills that are of an applied nature, rural areas have socio-cognitive skills that are geared more toward training as well as farming activity. Thus, it may be the case that rural areas are training labor which is then put to use in the most populous areas. In any case, as we have found farm activity contributes notably to the socio-cognitive lobe of the skills space in the most rural counties, such rural occupations may be a source of innovations and should not be overlooked by either planners or scholars of innovation. Despite such possibilities, further analysis examining temporal dynamics is needed to answer such questions.

Abbreviations

BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
CBSA	Core-Based Statistical Areas
ERS	Economic Research Service
IBRC	Indiana Business Research Center
MSA	Metropolitan Statistical Area
NAICS	North American Industry Classification System
NECTA	New England City and Town Areas
OMB	U.S. Office of Management and Budget
SOC	Standard Occupation Classification

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42854-022-00044-6.

Additional file 1. SOM Figure 1. Interdependence Values Calculated Using Different Geographies. **SOM Figure 2.** LMAs and CBSAs by Population Decile and Tightness. **SOM Figure 3.** Skills Network Dual Lobed Structure at County, LMA, and CBSA. **SOM Table 1.** Socio-Cognitive Specialty Skills of Counties in the Bottom Decile. **SOM Table 2.** Socio-Cognitive Skills of Counties in the Top Decile. **SOM Table 3.** Occupational Contribution of Socio-Cognitive Skills for Rural and Urban Counties.

Acknowledgements
The authors would like to acknowledge Tim Slaper, Riley Zipper, and Ping (Claire) Zheng for data support and discussion.

Authors' 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.

Availability of data and materials
We do not have permission to reproduce the raw employment data but it may be requested by researchers from the Indiana Business Research Center. All other data (O*NET) is publicly available as described in the text.

Declarations

Competing interests
The author(s) declare(s) that they have no competing interests.

Received: 18 August 2021 Accepted: 5 October 2022
Published online: 17 November 2022

References

Alabdulkareem A, Frank MR, Sun L, et al. Unpacking the polarization of workplace skills. Sci Adv. 2018;4:eaao6030. https://doi.org/10.1126/sciadv.aao6030.
Arthur WB. Foundations of complexity economics. Nat Rev Phys. 2021;3:136–45. https://doi.org/10.1038/s42254-020-00273-3.
Boschma R, Balland P-A, Kogler DF. Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. Ind Corp Chang. 2015;24:223–50. https://doi.org/10.1093/icc/dtu012.
Cavalletti B, Corsi M. "Beyond GDP" effects on national subjective well-being of OECD countries. Soc Indic Res. 2018;136:931–66. https://doi.org/10.1007/s11205-016-1477-0.
Elmqvist T, Andersson E, Frantzeskaki N, et al. Sustainability and resilience for transformation in the urban century. Nat Sustain. 2019;2:267–73. https://doi.org/10.1038/s41893-019-0250-1.
Fowler CS, Jensen L. Bridging the gap between geographic concept and the data we have: the case of labor markets in the USA. Environ Plan A. 2020;52:1395–414. https://doi.org/10.1177/0308518X20906154.

Gómez-Herrera E. Comparing alternative methods to estimate gravity models of bilateral trade. Empirical Economics. 2013;44:1087–111. https://doi.org/10.1007/s00181-012-0576-2.
Gotelli NJ. Null model analysis of species co-occurrence patterns. Ecology. 2000;81:2606–21. https://doi.org/10.2307/177478.
Griffith DM, Veech JA, Marsh CJ. Cooccur: probabilistic species co-occurrence analysis in R. J Stat Soft. 2016;69:1–17. https://doi.org/10.18637/jss.v069.c02.
Guevara MR, Hartmann D, Aristarán M, et al. The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations. Scientometrics. 2016;109:1695–709. https://doi.org/10.1007/s11192-016-2125-9.
Hidalgo CA. Economic complexity theory and applications. Nature Reviews Physics. 2021;3:92–113. https://doi.org/10.1038/s42254-020-00275-1.
Hidalgo CA, Klinger B, Barabási A-L, Hausmann R. The product space conditions the development of nations. Science. 2007;317:482–7. https://doi.org/10.1126/science.1144581.
Hidalgo CA, Balland P-A, Boschma R, et al (2018) The principle of relatedness. Springer, pp 451–457
Kogler DF, Rigby DL, Tucker I. Mapping knowledge space and technological relatedness in US cities. Eur Plan Stud. 2013;21:1374–91. https://doi.org/10.1080/09654313.2012.755832.
Liu G, Wong L, Chua HN. Complex discovery from weighted PPI networks. Bioinformatics. 2009;25:1891–7. https://doi.org/10.1093/bioinformatics/btp311.
Martin R, Sunley P. On the notion of regional economic resilience: conceptualization and explanation. J Econ Geogr. 2015;15:1–42.
Muneepeerakul R, Lobo J, Shutters ST, et al. Urban economies and occupation space: can they get "There" from "Here"? PLoS ONE. 2013;8:e73676. https://doi.org/10.1371/journal.pone.0073676.
National Center for O*NET Development (2020) O*NET database version 24.2
Neffke F, Henning M. Skill relatedness and firm diversification. Strateg Manag J. 2013;34:297–316. https://doi.org/10.1002/smj.2014.
Neffke F, Henning M, Boschma R. How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. Econ Geogr. 2011;87:237–65. https://doi.org/10.1111/j.1944-8287.2011.01121.x.
OECD. OECD guidelines on measuring subjective well-being. OECD. 2013.
Rigby DL. Technological relatedness and knowledge space: entry and exit of US cities from patent classes. Region Stud. 2015;49:1922–37. https://doi.org/10.1080/00343404.2013.854878.
Shutters ST, Waters K. Inferring networks of interdependent labor skills to illuminate urban economic structure. Entropy. 2020b;22:1078. https://doi.org/10.3390/e22101078.
Shutters ST, Muneepeerakul R, Lobo J. How hard is it for urban economies to become 'green'? Environ Plann B Plann Des. 2015a;43:198–209. https://doi.org/10.1177/0265813515600108.
Shutters ST, Muneepeerakul R, Lobo J. Quantifying urban economic resilience through labour force interdependence. Palgrave Communications. 2015b;1:15010. https://doi.org/10.1057/palcomms.2015.10.
Shutters ST, Muneepeerakul R, Lobo J. Constrained pathways to a creative urban economy. Urban Studies. 2016;53:3439–54. https://doi.org/10.1177/0042098015616892.
Shutters ST, Lobo J, Muneepeerakul R, et al. Urban occupational structures as information networks: the effect on network density of increasing number of occupations. PLoS ONE. 2018;13:e0196915. https://doi.org/10.1371/journal.pone.0196915.
Shutters ST, Waters K (2020a) Industrial structure and a tradeoff between productivity and resilience. Available at SSRN 3702545
Slaper T (2019) Rural economic development, wicked problems and complex adaptive systems. Indiana Business Review 94:
Tokuyama T. Algorithms and computation. Sendai, Japan: Springer; 2007.
U.S. Bureau of Labor Statistics. Occupational Employment and Wage Statistics (OEWS). 2018.
U.S. Bureau of Labor Statistics. Quarterly Census of Employment and Wages (QCEW). 2022.
U.S. Office of Management and Budget (2020) OMB Bulletin No. 20–01: Revised Delineations of Metropolitan Statistical Areas, Micropolitan Statistical Areas, and Combined Statistical Areas, and Guidance on Uses of the Delineations of These Areas. Washington, DC
Veech JA. A probabilistic model for analysing species co-occurrence. Glob Ecol Biogeogr. 2013;22:252–60. https://doi.org/10.1111/j.1466-8238.2012.00789.x.
Waters K, Shutters ST. Skills-approximate occupations: using networks to guide jobs retraining. Applied Network Science. 2022;7:43. https://doi.org/10.1007/s41109-022-00487-7.
West G. Scale: The universal laws of life, growth, and death in organisms, cities, and companies. Penguin; 2017.
Zheng P. Filling in the holes: Generating point estimates for QCEW suppressed data. 2020.

Publisher's Note