DEFINING AND MEASURING RURALITY IN THE US: FROM TYPOLOGIES TO CONTINUOUS INDICES

Brigitte Waldorf and Ayoung Kim

Department of Agricultural Economics, Purdue University
West Lafayette, IN 47907, USA
Email: bwaldorf@purdue.edu and kim1426@purdue.edu

Abstract. The paper provides an overview of quantitative measures of "rural" and "rurality." A distinction is made between discrete measures derived from threshold-based typologies, and continuous measures in the form of indices. We present a selection of classifications developed and employed in the United States. Subsequently the classifications are juxtaposed with an example of a continuous measure of rurality, the Index of Relative Rurality.

Key Words: rurality, rural classification, rural index

JRL Codes: C43, R58

Commissioned paper:
Workshop on Rationalizing Rural Area Classifications
[April 16/17, 2015: Keck Center, Washington, DC]

1. Introduction

Proper measurement of rurality is essential for researchers and policy makers alike. Researchers often try to assess whether rurality is a contributing factor to the variation in outcome variables, whether measured at a micro-scale such as individuals' health outcomes, or measured at the macroscale such as regions' population change, crime rate, or sustainability. For policy makers, proper measurement of rurality is of utmost importance as, for example, "needs analyses using inappropriate coding schemas can result in lack of funding or the implementation of policies that are ineffective for the targeted population" (Atav and Darling 2012, p.30).

This paper provides an overview of definitions and measures of "rural" and "rurality" developed and employed in the United States. It is comprised of six sections. Following the introduction, we begin with a general discussion on methods to measure, rank and code rurality in the second section. An important distinction is made between discrete measures based on typologies, and continuous measures in the form of indices. In the third section, we review a selection of rural-urban classifications that have been developed and employed in the United States. In the fourth section, we present the Index of Relative Rurality as an example of a continuous measure of rurality, and discuss its advantages over traditional typologies and definitions. In the fifth section, we exemplify the advantages of the Index of Relative Rurality using 2000 and 2010 data. The final section provides a summary and conclusions.

2. Methods for Measuring, Ranking, and Coding Rurality

2.1 Discrete Measures

The traditional approach to capturing rurality is the design of a typology or classification system for places or regions. In general, a typology is made up of three components: (a) a set of *n* objects to be assigned to discrete types; (b) the number of types, *k*; (c) the criteria governing the assignment of objects to types.

Objects: in the case of delineating rurality, the objects are areas such as counties, census tracts, grid cells or, in the extreme, points. The choice of objects ultimately determines the scale at which research on rurality can be conducted, as well as the scale at which policies can be implemented. Typically, the internal heterogeneity of spatial objects rises with increasing spatial scale. However, a very small spatial scale becomes impractical if data are unavailable at such small spatial scales. Choosing the proper scale is tightly related to what Cromartie and Bucholz (2008) refer to as the challenge of choosing an appropriate urban boundary.

Number of types: The simplest typology is a dichotomy that assigns *n* objects to *k*=2 types, such as rural versus non-rural. While such a binary typology might be appealing because of its simplicity, it is also unsatisfactory because it does not do justice to the complexity and diversity of both the rural and non-rural landscapes. A larger number of rural types allows policy makers to, for example, target resource allocations more efficiently.

Criteria: The selection of criteria assigning objects to types is at the very core of any classification as the criteria speak to the underlying dimensions of rurality as well as the mechanism of assigning objects to types. Due to the elusive character of rurality, different dimensions have been utilized in the various rural-urban classifications. The most commonly addressed dimensions are size and density. Researchers agree that—ceteris paribus—places with small populations are more rural than places with large populations, and places with low density are more rural than places with high density. Remoteness is a less frequently used dimension in existing typologies, possibly because it is more difficult to measure. The remoteness dimension is based on the idea that remote places are more rural than less remote places. Other dimensions used to capture rurality include commuter linkages, land use, and economic function.

In addition to identifying the underlying dimensions of rurality, the criteria also determine the mechanism by which an object is assigned to a particular type. For most rural-urban classifications, the assignment is based on thresholds that separate the various types. We typically use "ball park figures" for the threshold, such as "500 persons per square mile" or "20,000 residents." Such thresholds are arbitrary and reflect our preference for "round numbers." That is, we have yet to see a rural-urban classification using a size threshold of 483,278 persons.

The thresholds are also not context specific. For example, when defining urban places, the U.S. Census Bureau uses the same thresholds for Alaska as it does for New Jersey. Moreover, since the objects to be classified are spatial units, for instance counties or census districts, threshold-based categorization are not independent of the spatial scale. New thresholds need to be selected and justified when using a different spatial scale.

Finally, thresholds create "artificial" similarities and dissimilarities. For example, take a dichotomous categorization based on just one variable and one threshold – say greater or smaller than 500. For three objects—A, B, and C with values 2, 499, and 501, respectively—the quite dissimilar objects A and B will be assigned to the same type, whereas the rather similar objects B and C are assigned to different types.

As an alternative to using thresholds as the assignment mechanism, similarity measures can be used to group n object to k types using m variables. Similarity is measured as distance in m-dimensional space, and the assignment of objects to types is based on maximizing the similarity of objects within a type. Such techniques—also referred to as cluster analyses—differ depending on the kind of distance measures used and the specifics of the assignment protocol (for example, hierarchical versus non-hierarchical).[1]

Whether based on thresholds or on similarity measures, m-dimensional typologies do not include mechanism that allow us to rank the k types along a continuum from most urban to most rural. Such assignments are made subjectively by an implicit ordering of the relative importance of the underlying dimensions. For example, suppose nonmetropolitan counties are classified into four groups, using two criteria – size (more vs. less than 20,000 inhabitants) and adjacency to metro area (adjacent or not adjacent). There will undoubtedly be agreement that the large counties with more than 20,000 inhabitants adjacent to a metro area are the most urban counties, whereas small counties not adjacent to metro areas are the most rural. However, there is no "natural ordering" of the two "in-between" types and we may disagree as to whether small counties adjacent to a metro area are more or less urban than big counties not adjacent to a metro area.

[1] Appendix A provides an example of such a classification of US counties based on four criteria. Zhou et al. (2004) used a cluster analysis-based classification for their air quality research in California, yielding four types for 7 variables and 1,895 census tracts.

The same problem arises with similarity-based typologies. The outcome of such a classification are *k* groups of spatial units of which we know that they are quite similar. However, without further inspection and subjective decisions, similarity-based typologies do not include a ranking along a rural-urban continuum.

2.2 Continuous Measures

Most certainly, rurality is not the only concept that is difficult to quantify. Quantifying vulnerability, resilience, sustainability and development poses similar difficulties (Pillarisetti and van den Bergh 2010). These problems are foremost due to the concepts' vague definitions and multidimensionality. A key question is how can we measure and combine these different dimensions and end up with a one-dimensional measuring stick that allows us to compare objects by their degree of rurality (or vulnerability or development). A very successful and elegant approach to tackle the problem has been implemented by the United Nations. Specifically, the United Nations Development Programme (UNDP 2014) uses a continuous aggregate index to measure countries' levels of development with the Human Development Index (HDI). The HDI is a continuous, threshold-free measure that is responsive to even slight changes in any of its three dimensions (health, knowledge, and standard of living).[2] Waldorf (2006, 2007a, 2007b) similarly designed a continuous, threshold-free index of rurality, the Index of Relative Rurality (IRR) that ranges between 0 (most urban) to 1 (most rural).

The design of the IRR is similar to the HDI design and consists of four steps: (1) identifying the dimensions of rurality; (2) selecting measurable variables to adequately represent each dimension; (3) re-scaling the variables onto comparable scales; (4) selecting a function that links the re-scaled variables so that multidimensionality is reduced into one-dimensionality, i.e., $f(\cdot):\mathbb{R}^n \rightarrow \mathbb{R}^1$. Each step of the design involves one or more subjective decisions that are, however, made explicit and can be based on defendable justifications. It should be kept in mind, though, that – due to the elusive nature of the rurality concept – it will ultimately be impossible to assess the "precision" of the measure.

3. Overview of Rural-Urban Classification Systems Used in the US

In the US, a large number of typologies have been developed to operationally identify rural places. They differ with respect to one or more of the three essential components, i.e., objects, types, and criteria. We begin the overview with the rural-urban classification system of the U.S. Census Bureau since it is unique in its reliance on a very small spatial scale and since it is used in other typologies as a criterion to assign places to particular types. Moreover, the rural-urban system of the U.S. Census Bureau is not only a typology of tracts into the rural-urban dichotomy, but also a delineation of urban areas. Second, we will discuss the Core Based Statistical Area classification and delineation of the Office of Management and Budget (OMB) because it too forms a building block of many other rural-urban typologies. In fact, in the U.S., the vast majority of typologies are derivatives of, or utilize one or both of these foundations.

---

[2] In its most basic form, the HDI is a multidimensional measure on a continuous scale from 0 to 100 capturing three dimensions: (1) long and healthy life; (2) knowledge; and (3) a decent standard of living.

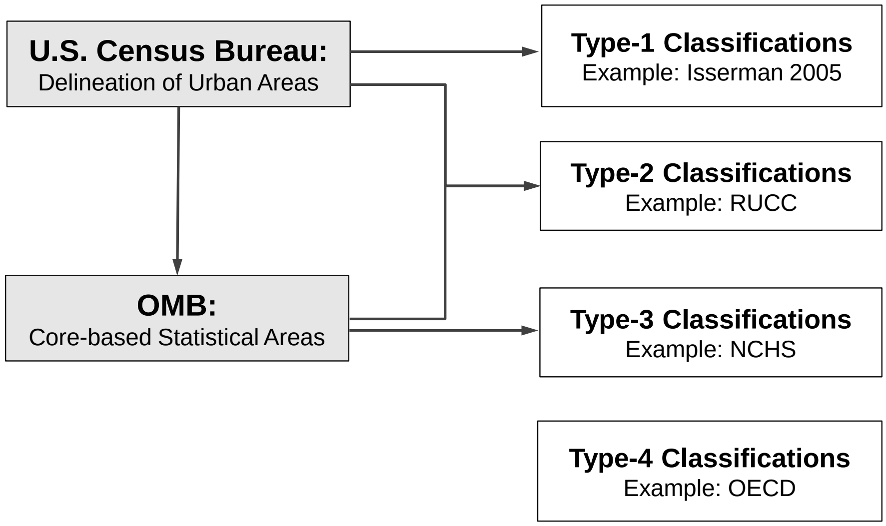


Figure 1. Connections among Rural-Urban Classification Systems

Figure 1 shows the reliance of rural-urban classifications for the United States on one or both of these building blocks from the U.S. Census Bureau and the OMB. We can distinguish four types. That is, classifications that rely on:

- urban areas of the Census Bureau (Type-1. For example: Isserman (2005));
- urban areas of the Census Bureau and OMB definitions (Type-2. For example: ERS classifications);
- OMB definitions (Type-3. Example: NCHS classification);
- neither Census Bureau nor OMB classifications (Type-4. For example: OECD typology).

Subsequent to describing the classifications of the U.S. Census Bureau and the OMB, we review a selection of the classifications belonging to the four types. The selected classifications are summarized in Appendix B.

3.1 The Rural-Urban Typology of the US Census Bureau.

Population size, population density and contiguity are the three main criteria used to delineate urban areas. An urban area is defined as a contiguous area that has, in its core, a population density of at least 1,000 persons per square mile and that has a total population of 2,500 or more residents.[3] For the 2010 definition, the smallest spatial unit is the census tract. Using a population size threshold, the Census Bureau further distinguishes two types of urban areas: (1) urbanized areas have at least 50,000 residents; and (2) urban clusters that have fewer than 50,000 residents. The territory outside of urban areas is defined as rural. The population is labelled "rural" or "urban" depending on the residence location being inside or outside an urban area. The percent of the US population living in rural areas was 24.8% in 1990, 21% in 2000, and 19.3% in 2010.[4]

[3] The delineation of urban areas is quite complex. Detailed definitions and criteria for the 2000 and 2010 urban and rural classifications are provided at https://www.census.gov/geo/reference/ua/urban-rural-2000.html and https://www.census.gov/geo/reference/ua/urban-rural-2010.html, respectively.
[4] Note that the Census Bureau's definitions of what is "urban" changed over time.

3.2 OMB's Core Based Statistical Areas

OMB delineates the so-called Core Based Statistical Areas (CBSA) (see http://www.whitehouse.gov/omb/bulletins/b03-04.html). A CBSA consists of one or more counties that jointly form a *contiguous* area. Two types of counties within a CBSA are distinguished. (1) Central counties are counties in which at least 50% of the population lives in an urban area of 10,000 residents or more. Every CBSA must have at least one central county. (2) Outlying counties have strong commuter ties with the central counties (at least 25% of the employed residents in an outlying county must work in the central county (counties), or at least 25% of the outlying county's labor force must reside in the central county (counties). Core Based Statistical Areas are named after their principal city (cities). Principle cities include the largest city of the CBSA plus additional cities that meet specified size criteria.

Two types of CBSAs are distinguished. First, a CBSA that includes an urban area with at least 50,000 residents is defined as Metropolitan Statistical Area (MSA). Second, a CBSA that includes an urban area with at least 10,000 urban residents but fewer than 50,000 is labeled Micropolitan Statistical Areas (MiSA). Counties not belonging to either a metropolitan or a micropolitan statistical area are referred to as "Noncore" counties.

Noteworthy is the distribution of urban and rural populations (as defined by the US Census Bureau) across the three types of counties. The Noncore counties are not entirely composed of rural residents. In fact, in 2010, 25% of residents living in Noncore Counties were classified as "urban" according to the US Census Bureau. And, a good deal of the population in CBSAs was classified as rural: 47.1% in micropolitan counties, and 11.7% in metropolitan counties. This seeming contradiction is due to the commuter flow criterion used for CBSA delineation. On the one hand, a CBSA includes—by design—primarily rural counties that are functionally linked with the highly urbanized central counties via commuter flows. On the other hand, counties with up to 9,999 urban residents are classified as Noncore county as long as they are not functionally connected to a CBSA via commuter flows.

OMB emphasizes that the CBSA classification is different from a rural-urban classification, stating explicitly: "The CBSA classification does not equate to an urban-rural classification; Metropolitan and Micropolitan Statistical Areas and many counties outside CBSAs contain both urban and rural populations." (Office of Management and Budget 2000, p. 82236). Yet, as Fuguitt (2005) claims, the simple metro/non-metro distinction may be an appropriate starting point for policy-oriented research, and countless scholarly articles used the metro/non-metro distinction as a proxy for an urban/rural distinction. (e.g., Domina 2006, Glasgow and Brown 2012, Johnson and Fuguitt 2000, Kandel et al. 2011, Levernier et al. 2000, Renkow and Hoover 2000, Waldorf et al. 2008, Weber 2007).

3.3 Type-1 Classifications

3.3.1 Isserman (2005).

Isserman's 'Rural-Urban Density Typology' (Table 1) utilizes thresholds for four variables – percentage of urban residents; total number of urban residents; population density; and population size of the county's largest urban area – to define 1,790 rural, 1,022 mixed rural, 158 mixed urban, and 171 urban counties in the US. Isserman's typology deliberately avoids the—according to him—misleading metro/non-metro classification. As a result, his classification undoubtedly does a good job at identifying the extremes. That is, the "urban status" of urban counties is unquestioned[5] and so is the "rural status" of counties labeled "rural." However, the typology does a less satisfactory job in separating the two mixed categories. In fact, the group of counties that do not meet either the rural or the urban thresholds, are only differentiated by a population density threshold of above versus below 320 persons per square-mile.

Table 1. Isserman's Rural-Urban Density Typology of U.S. Counties

		Population density [pers./mile^2]	% urban	Population size of largest urban area[a]	Total number of urban residents
Extremes	Rural	<500	< 10%	< 10,000	
	Urban	500+	90% +		50,000 +
Counties not meeting the rural or urban thresholds:					
Mixed	Mixed Rural	<320			
	Mixed Urban	320+			

[a] As defined by the U.S. Census Bureau.

3.3.2 Veterans Health Administration (VHA)

The VHA rural-urban classification system is an example of a system that emphasizes a nuanced rural differentiation. VHA developed a three-tier classification in which locations are categorized as urban, rural, or highly rural (Berke et al. 2009). A location is classified as urban if it is inside an urbanized area (as defined by the US Census Bureau). A location is classified as rural if it is located in a county with a population density of at least 7 persons per square mile, but outside an urbanized area. Finally, counties with a population density of less than 7 persons per square mile are categorized as highly rural. Note that this system is also unique in that it uses a merger of county boundaries and boundaries of urbanized areas to delineate the three territory. Note also that urban clusters are considered rural or even highly rural. Most recently, VHA and their Office of Rural Health announced to abandon the system and adopt the Rural-Urban Commuting Area (RUCA) system instead (http://www.ruralhealth.va.gov/rural-definition.asp).

3.4 Type-2 Classifications

[5] Isserman noted that some urban counties include a substantial portion of undeveloped land or farmland.

3.4.1 Rural-Urban Continuum Code

ERS has proposed two important typologies, The Rural-urban Continuum Code (RUCC) and the Urban Influence Code (UIC). The RUCC typology is a refinement of the tri-partite classification of counties into Metropolitan, Micropolitan and Noncore counties. The RUCC typology is hierarchical with the first level separating metropolitan from non-metropolitan counties. At the second level, metropolitan counties are differentiated into three groups (RUCC 1 to 3), using MSA size as the distinctive criterion, whereas non-metropolitan counties are differentiated into six groups (RUCC 4 to 9) using urban population size (as defined by the U.S. Census) and adjacency to a metropolitan area as distinguishing criteria. Increasing numbers are meant to reflect increasing rurality. The name (Rural-Urban *Continuum* Code) as well as the numeric coding suggest a "continuous" and monotonic increase of rurality on a nine-point scale. However, this suggestion may actually be a dangerous deception as it hides the initial distinction between OMB's metro (codes 1 to 3) and non-metro counties (codes 4 to 9), a distinction that is not intended to mirror a classification by counties' degree of rurality as emphasized by the OMB.

3.4.2 National Center for Education Statistics (NCES)

Whereas many typologies classify counties, other agencies may need to develop their own typology because counties are not the most appropriate spatial scale for their policies or research. NCES, for example, used to classify school locations with the RUCC system, but now uses longitude and latitude information to classify places/points/school locales into four main types:

- City – located within a principal city (as designated in the CBSA system);
- Suburb – located outside a principal city, inside an urbanized area (as defined by the US Census Bureau);
- Town – located inside an urban cluster;
- Rural – located inside rural territory as defined by the US Census Bureau.

Each type is further differentiated into three subtypes, using size and—in the case of schools in town and rural locales—distance criteria.

3.4.3 Department of Housing and Urban Development (HUD)

HUD has special programs to assist and support housing in rural areas. (http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/economicdevelopment/programs/rhed). Toward that end, HUD uses size criteria in combination with location inside/outside a metro area to define rural. Rural are all places with fewer than 2,500 residents and all counties/parishes with fewer than 20,000 urban residents (as defined by the U.S. Census Bureau). Places outside a metro area are rural if they have fewer than 20,000 residents

3.5 Type-3 Classifications

3.5.1 Rural Metropolitan Interface Levels

The definition of Rural-Metropolitan Interface Levels (Waldorf 2007) uses the Index of Relative Rurality[6] (Waldorf 2006) to develop a traditional threshold-based classification system of US counties. In total, seven levels are distinguished that speak to differences in the way rurality plays out in places within the influence of a metropolitan core versus in places far away from a metropolitan core. Reasons for these differences include, for example, accessibility to the amenities of urban centers as well as the spillovers of agglomeration economies. The assignment of counties to the seven Rural-Metropolitan Interface Levels is a two-step procedure. In the first step, counties are assigned to one of three spheres: metropolitan, interface, rural. In the second step, counties within a spheres are further differentiated based on size, location relative to a metropolitan area, and degree of rurality. In total, seven different types (levels) are distinguished. Table 2 summarizes the classification system.

Table 2. *Rural Metropolitan Interface Levels*

		Location Relative to Metro Area	Degree of Rurality
Metropolitan Sphere			
A	metropolitan central county with 500,000+ residents.	Within	Low
B	metropolitan central counties with < 500,000 residents.	Within	Low.
C	outlying metropolitan counties with an IRR < 0.4	Within	Low
Rural-Metropolitan Interface			
D	outlying metropolitan counties with an IRR ≥ 0.4	Within	High
E	nonmetropolitan counties adjacent to a metro area, IRR < 0.4	Adjacent	Low
F	nonmetropolitan counties adjacent to a metro area, IRR ≥ 0.4	Adjacent	High
Rural Sphere			
G	nonmetropolitan counties not adjacent to a metropolitan area	Remote	High

The metropolitan sphere includes metropolitan counties with a low degree of rurality. Within this sphere, Level A counties are metropolitan central counties with at least half a million residents; Level B counties are metropolitan central counties with less than half a million residents; and Level C counties are outlying metropolitan counties with an IRR < 0.4; they are often suburban in character. The interface sphere includes counties that combine high accessibility and high rurality, or have medium accessibility. Three levels are distinguished within this sphere: Level D counties are outlying metropolitan counties with an IRR ≥ 0.4; Level E counties are non-metropolitan counties adjacent to a metro area with IRR < 0.4; Level F counties are non-metropolitan counties adjacent to a metro area with IRR ≥ 0.4. Finally, the rural sphere is not further differentiated and includes Level G counties, defined as nonmetropolitan counties not adjacent to a metropolitan area.

[6] See section 4.

3.5.2 Urban Influence Code

Just like the RUCC, the Urban Influence Code (UIC) is also a refinement of the metro / nonmetro dichotomy. The criticism voiced against the RUCC thus similarly applies to the UIC. The UIC is a hierarchical typology with metro and nonmetro counties being separated at the first level. Subsequently metro counties are assigned to two types depending on whether the metro area has at least a million residents (UIC 1) or not (UIC 2). Nonmetro counties are assigned to 10 types (UIC 3 to 12) based on type of nonmetro county (micropolitan, noncore county with town of 2,500, noncore county without town of 2,500 residents); adjacency to large metro area, adjacency to small metro area, not adjacent to metro area. On the scale from 1 to 12, increasing numbers are meant to reflect a decreasing urban influence.

3.5.3 National Centers for Health Statistics (NCHS)

Like the ERS classifications, the urban-rural classification system of NCHS is also based on a metro/nonmetro distinction of counties. However, it is different from the ERS classifications in that it is biased towards a finer urban differentiation for very large metro areas. The system emphasizes the urban differentiation so as to provide a nuanced look at rural-urban health differences which are primarily manifested in large rural-suburban mortality and morbidity disparities.

The NCHS uses three criteria: CBHS status, size, and a core/fringe distinction within MSAs of more than a million inhabitants (Ingram and Franco 2014). The rural (non-metropolitan) regime remains quite undifferentiated, only distinguishing between micropolitan and non-core counties.

3.6 Type-4 Classifications

3.6.1 OCED Typologies

The OECD typology is geared towards assigning regions to one of three types: *predominantly urban*, *intermediate*, or *predominantly rural* (OECD 2011). Towards that end, the smaller areal units (counties in the case of the US) within the regions are classified as rural or urban, The classification of rural versus urban counties is entirely based on population density, choosing a threshold of 150 persons per square-kilometer (389 persons per square mile). Regions are subsequently assigned to one of the three types based on the share of rural counties in the region, and the size of urban areas. An extended OECD regional typology takes proximity to urban centers into account and distinguishes five types: predominantly urban, intermediate close to city, intermediate remote, predominantly rural close to city, and predominantly rural remote. The threshold for city size is 50,000 residents. The threshold for proximity is half the inhabitants being 60 driving minutes away from the city (Brezzi et al. 2011).

3.6.2 Ayres et al. (2012) Typology

Ayres et al. (2012) designed a county-based typology specifically designed for the state of Indiana. It utilizes dimensions suggested by Isserman (2005)—i.e. county population, population density—plus the size of the largest city or town in the county. The thresholds are chosen so as to specifically fit the situation in Indiana. Moreover, also included is a dimension called "county identity." County identity is a subjective indicator reflecting whether the residents perceive their county as rural.

4. The Index of Relative Rurality

The Index of Relative Rurality has not yet been applied in a policy context. But, it has found its way into research from a broad range of disciplines (Hubach et al. 2014, Barber 2013, Kaza 2013, De Montis et al. 2012, Gallardo and Scammahorn 2012, Heflin and Miller 2012, Mammen et al. 2011, Lambert et al. 2010, Stewart and Lambert 2008). This section begins by detailing its design, followed by a discussion of its advantages relative to typology-based rurality measures.

4.1 Design

The Index of Relative Rurality (IRR) is an aggregate index, designed by Waldorf (2006) as an alternative approach to discrete threshold-based classifications. As outlined in section 2, the design of the IRR is similar to the HDI design and consists of four steps.

Step 1: Identifying the dimensions of rurality. Four dimensions of rurality enter the index: size, density, remoteness, and built-up area. The first two dimensions, size and density, are represented in almost all existing rural typologies (see Appendix B). Thus, there seems to be agreement that, ceteris paribus, places with small populations are more rural than places with large populations, places with low density are more rural than places with high density. Remoteness is a less frequently used dimension in existing typologies, perhaps because it is more difficult to measure. It is based on the idea that remote places are more rural than less remote places. The last dimension, built-up area, captures the idea that places with few built-up areas are more rural than heavily built-up places.

The question arises whether there are additional dimensions of rurality that should be included in the design.[7] In the past, it may have been defendable to include the reliance on agriculture as a key dimension. However, today's agriculture accounts for such a small share of economic activities even in areas traditionally perceived as rural, that it no longer qualifies as a key dimension. Similarly, many socio-economic characteristics often associated with rural areas—lower wages, for example—are often outcomes rather than defining dimensions of rurality.

Step 2. The selection of variables. Choosing variables that can adequately represent each dimension is of course very much dependent on data availability. In the original 2006 design of the index,[8] simple measures were chosen that can be easily replicated, updated, and are available at different spatial scales. The size dimension is captured via the logarithm of the population size and the density dimension is operationalized as the logarithm of population density. The logarithmic transformations for population size and density were chosen to correct for their skewed distributions (abundance of places with small populations and low densities, whereas places with large populations and high densities are rare). For the remoteness dimension and the built-up area dimension, the variables used in original index were updated for the presentation in section 5. Spherical

[7] Note that the number of dimensions that can be incorporated in the index is flexible.

[8] The original version of the IRR was introduced in a conference paper by Waldorf (2006). Portions of this paper are similar to the 2006 conference paper.

distance to the midpoint of the closest central county of a metro area was replaced with network distance;[9] the percentage of the population living in an urban area was replaced the urban area as a percentage of total land area in the updated version

Step 3: Rescaling of variables. The variables are measured on different scales. Distance, for example, is measured in miles whereas the built-up area variable is measured in percentages. Consequently, a re-scaling is needed to ensure scale compatibility across the four dimensions as well as independence from the initial units of measurement (kilometers or miles, for example). One possibility for re-scaling is to express the observed values in standard deviations from the mean. The standard deviation scale ranges from minus infinity to plus infinity. Another possibility is a bounded scale, for example one that ranges from "0" to "1". IRR is based on the bounded scale where "0" represents the most urban object and one represents the most rural object. The re-scaling procedure involves a simple transformation:

- For variables that are negatively related to rurality, such as size, density and built-up area, the rescaling of an observed value of the variable X, X_i, the re-scaling takes on the form:

$$X_i \rightarrow \frac{X_{\max} - X_i}{X_{\max} - X_{\min}} \in [0,1]$$

  where X_{min} and X_{max} are the minimum and maximum values of X, respectively, observed across the spatial areas or places.
- For variables that are positively related to rurality, such as distance to a metro area, the rescaling takes on the form:

$$X_i \rightarrow 1 - \frac{X_{\max} - X_i}{X_{\max} - X_{\min}} \in [0,1]$$

Step 4: Selecting a link function. An important step is the selection of a link function. This function should reflect how the four dimensions jointly determine the rurality of a place. At issue is the relative importance of the four dimensions for rurality. For example, should density be given more weight than size and remoteness? In the absence of any theoretical guidance, the most simple link function was chosen, namely the unweighted average of the four rescaled variables. The resulting index – the Index of Relative Rurality – is a relative measure because it places the rurality of a spatial unit within the wider context of the rurality of all spatial units considered.

4.2 IRR's Advantages over Threshold-based Typologies

IRR has several advantages over existing typology-based rurality measures. First, aside from data availability constraints, the measure is not confined to a particular spatial scale. In its original specification (Waldorf 2006) it was designed for counties. However, the index-approach to capture degrees of rurality can also be applied to groups of counties, which increasingly form the basis of regional development efforts, as well as to smaller scales such as townships or census tracts. Muhlenkamp and Waldorf (2008), for example, adapt the index to Public Use Micro Areas. The scale-flexibility is an important advantage

[9] The network distance is the shortest route along the network of roads and highways.
http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=An%20overview%20of%20Network%20Analyst

over traditional classifications and will be particularly beneficial for designing and evaluating local policies and regional development strategies.

Second, the Index treats rurality as a relative concept. The underlying question is not whether a place is rural, but what is the place's rurality relative to other places. As such, the index takes into account the context which may change over space and over time.

Third, the Index of Relative Rurality is a continuous measure that is responsive to the multi-faceted nature of rurality and is sensitive to even small changes in one or several of the defining variables. It is thus well suited to investigate the trajectories of rurality over time. In contrast, when using threshold-based typologies, changes in a place's characteristics are only recorded as a change in status, say from mixed rural to mixed urban, if the changes are big enough to move beyond one or more thresholds.

Finally, as a continuous variable, the Index also opens up analytical possibilities. For example, the using the Index in regression models, one can conveniently test not only whether there is a rurality effect, but also whether the effect is nonlinear. With rurality measured via threshold-based classifications, such tests are either not possible or awkward as the rurality data are categorical and need to enter analytical models as a series of dummy variables.

5. Exploring the Index of Relative Rurality

In this section, we present the 2000 IRR and 2010 IRR using the updated variable selection[10] described in section 4.1, and demonstrate IRR's advantages relative to threshold-based rurality classifications. The data for size, density and built-up area were taken from the US Census Bureau Decennial Census 2000 and 2010. Data for network distances (remoteness dimension) were taken from the geospatial database of the U.S. Census Bureau and Federal Highway Administration.

5.1 Correspondence between IRR and the ERS Measures of Rurality

Following the example of Gelphi and Parker (1997), we demonstrate how the IRR compares to the well-known ERS classifications. Table 3 shows IRR summary statistics, calculated separately for all counties in each of the nine RUCC categories. The average IRRs mirror the RUCC categories quite well, in that higher RUCC code are, on average, associated with higher IRR values. For the 432 counties classified as a category-1 in the RUCC scheme, the average IRR index is 0.37. At the other extreme, for the 424 category-9 counties, the index averages to 0.62.

[10] Upon request, the authors will share the 2000 and 2010 IRR data for U.S. Counties.

Table 3. Comparison of RUCC and IRR

RUCC (2013)	# of Counties	Index of Relative Rurality 2010				
		Average	Stdev	Min	Max	Range
1	432	0.37	0.13	0.04	0.58	0.54
2	379	0.44	0.08	0.16	0.65	0.48
3	354	0.47	0.07	0.17	0.68	0.51
4	214	0.48	0.03	0.22	0.60	0.38
5	92	0.51	0.03	0.43	0.69	0.26
6	593	0.53	0.03	0.22	0.63	0.41
7	433	0.56	0.05	0.35	0.89	0.54
8	220	0.57	0.04	0.51	0.73	0.22
9	424	0.62	0.06	0.52	0.85	0.33
Grand Total	3,141	0.50	0.10	0.04	0.89	0.85

However, the within-category variations are quite big, especially for the top-three RUCC categories that collectively include all metropolitan counties. For example, for the 432 counties classified as category-1 counties in the RUCC typology (i.e., they are part of a metropolitan area with at least one million residents), the IRR ranges between 0.04 and 0.58, with an average of IRR = 0.37. The reason for the wide variation is that metropolitan areas also include small, sparsely populated counties—characteristics that lead to high IRR scores—with a sizable commuter flow into the core of the metro area—a characteristic not included in the Index.

Table 4. Comparison of UIC and IRR

UIC (2013)	# of Counties	Index of Relative Rurality 2010				
		Average	Stdev	Min	Max	Range
1	432	0.37	0.13	0.04	0.58	0.54
2	733	0.45	0.08	0.16	0.68	0.51
3	130	0.49	0.03	0.42	0.60	0.18
4	149	0.53	0.04	0.31	0.66	0.35
5	242	0.50	0.04	0.22	0.65	0.43
6	344	0.54	0.04	0.22	0.63	0.41
7	162	0.58	0.04	0.50	0.73	0.22
8	269	0.54	0.05	0.38	0.78	0.41
9	184	0.56	0.03	0.35	0.65	0.31
10	189	0.60	0.05	0.51	0.81	0.30
11	125	0.59	0.07	0.47	0.89	0.42
12	182	0.63	0.06	0.52	0.85	0.33
Grand Total	3,141	0.50	0.10	0.04	0.89	0.85

Table 4 reveals that the Urban Influence Code (UIC) does not correspond as well with the Index as the RUCC. Increasing UIC codes are not reflected in increasing average IRR values. In fact, the monotonic increase is disrupted at three points (UIC = 5, UIC = 8, and UIC = 11). Moreover, the IRR heterogeneity within the two metropolitan codes (UIC = 1 and UIC = 2) is quite substantial.

5.2 Flexible Spatial Scale

The Index of Relative Rurality is not confined to a particular spatial scale, such as counties. Instead, it can also be applied to groups of counties as well as to smaller scales such as townships or census tracts. Such smooth transfer from one spatial scale to another is not possible for the traditional threshold-based classification systems, as the thresholds need to be specified (and justified) separately for each spatial scale. The transfer capabilities are thus an important advantage over traditional classifications and will be particularly beneficial for designing and evaluating regional development strategies.

The scale flexibility is shown in Figure 4. Using the same four variables as for the nation-wide county measure, the maps show the IRR for Indiana census tracts on the left, and zip code areas on the right.

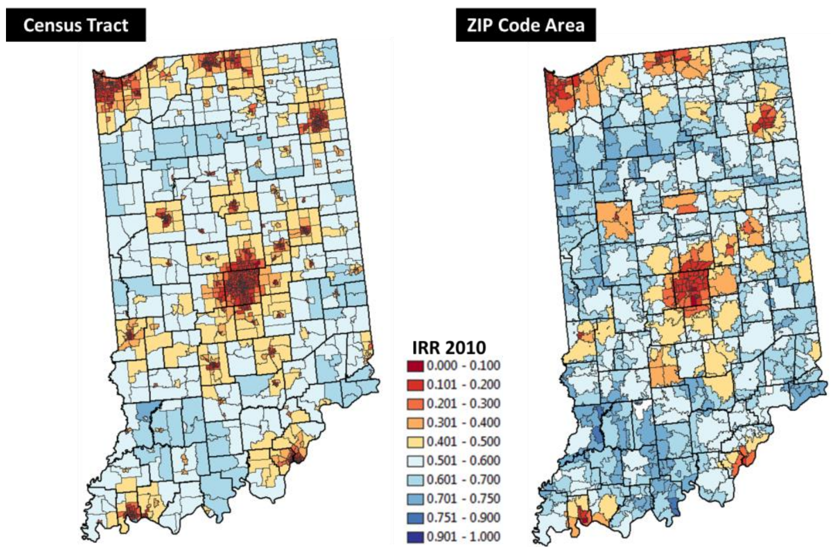


Figure 2. IRR for Indiana Census Tracts and Zip Code Areas, 2010

5.3 Rurality Variations across Space and Time

The Index of Relative Rurality is ideally suited to capture variations in rurality across space and over time. Figures 3 shows the Index for counties in the continental U.S. for the year 2000.[11] Not surprisingly, the lowest IRR scores (i.e., highly urbanized counties) are found along the coasts as well as around the urban centers along the Great Lakes. Particularly interesting is the upward trend in IRR scores as one moves from the Midwest to the Great Plains. In fact, counties east of the Mississippi tend to have low to medium levels of rurality while higher degrees of rurality (IRR>0.6) – so prevalent in the Great Plains – is almost absent. Extreme rurality, say IRR > 0.7, is entirely confined to the West. In fact, the top-10 most rural counties are all located in Alaska. The most rural counties outside of Alaska were Loving, TX (ranked 13) and Esmeralda, NV (ranked 18).

---

[11] See Appendix C for a map of the IRR in 2010, and its approximation via a 3rd-order polynomial.

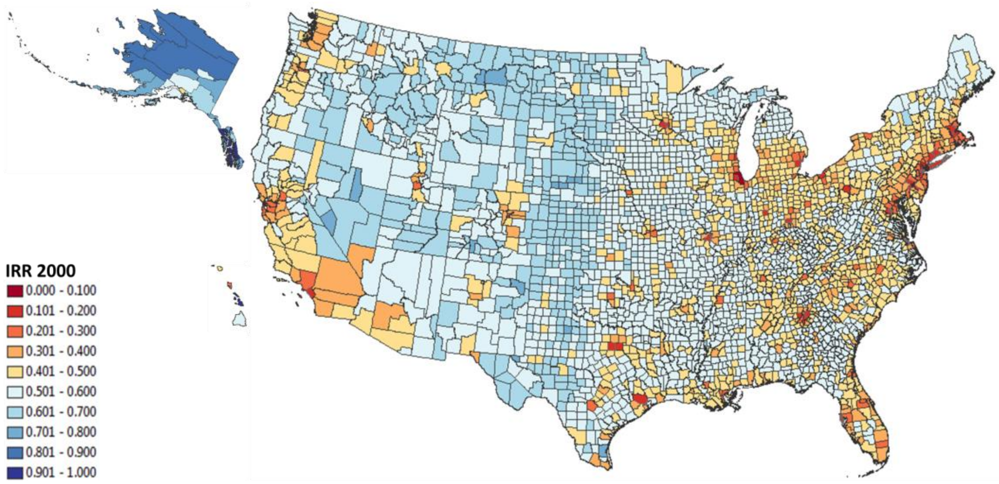


Figure 3. Index of Relative Rurality (IRR) for U.S. Counties, 2000

By 2010, small changes had occurred in the spatial pattern of rurality in the US. For example, in 2000, the top-10 most urban counties included eight counties of the East Coast Megalopolis, Cook County, IL (Chicago) and San Francisco, CA. Ten years later, Baltimore, rather than DC, made it into the top-10, and Suffolk County, MA (Boston) became slightly more urban, moving from rank 9 to rank 8. At the other extreme of the most rural counties, by 2010 Alaska had become even more rural relative to the rest of the country: 17 of the 29 counties in Alaska were more rural than any county outside of Alaska. Loving County, TX had moved from rank 13 in 2000 to rank 19 in 2010, and Esmeralda County had climbed from rank 18 to rank 22. Note that none of these changes precipitates into changes in, for example, RUCC or IUC coding.

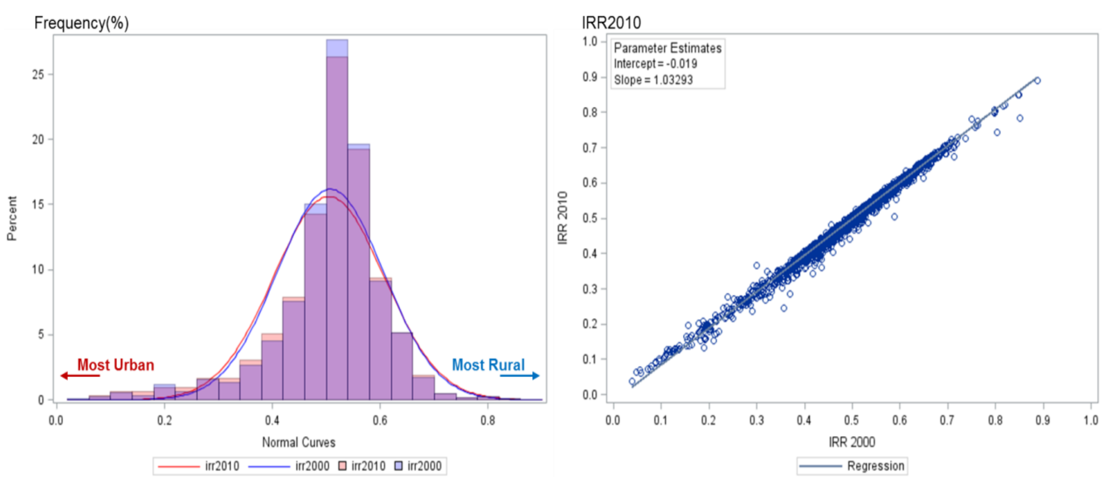


Figure 4. Temporal Persistence of Rurality in U.S. Counties
Left: Histogram of 2000 and 2010 IRR; Right: Scattergram of 2000 and 2010 IRR

Figure 4 shows that overall—including the wide spectrum of IRR scores away from the extremes—the pattern of rurality has barely changed during the first decade of the 21st

century. There was a slight shift towards more urbanization with the average IRR declining from 0.507 in 2000 to 0.505 in 2010. For the most part, however, the changes are small and few counties changed their relative standing. The scattergram of the IRR_{2000} versus IRR_{2010} convincingly shows this persistence, with most counties scattered around the regression line, which is almost indistinguishable from the 45-degree line.

The temporal persistence of the rurality pattern is not surprising, at least within the ten-year horizon portrayed here. Figure 5 shows that counties with decreasing IRR (becoming more urban) are concentrated in the western half of the United States, as well as along the entire East Coast from Boston to Florida. Going from the East coast to the Rocky Mountains, the incidence of counties becoming more urban decreases. Instead, counties that became more rural, are almost exclusively located East of the Rocky Mountains and concentrated in the Great Plains and South. These patterns reflect the ongoing urbanization and urban sprawl in the western U.S. as well as the de-population in some of the interior east of the Rocky Mountains.

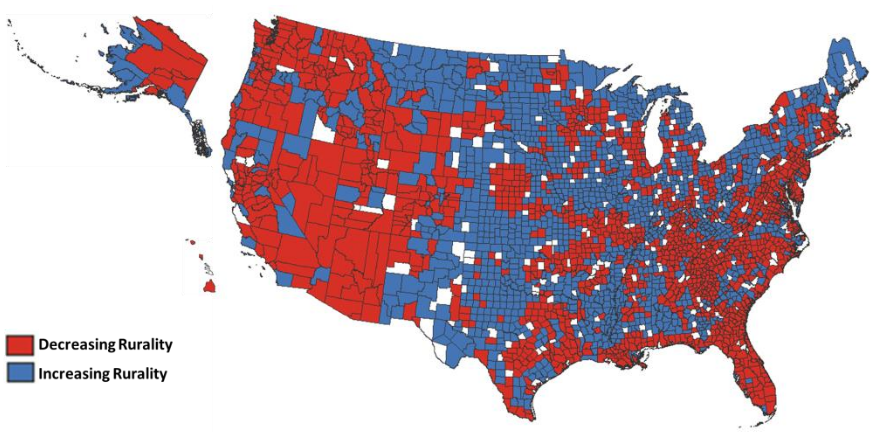


Figure 5. Spatial Pattern of Rurality Change in U.S. Counties between 2000 and 2010

This general pattern of urbanization differences between the coasts and the interior is also visualized in Figure 6. Estimated changes in rurality are expressed as a nonlinear function (2^{nd}-order polynomial) of longitude. The blue (red) line shows the regression of IRR on longitude for all counties north (south) of 38.3 Northern latitude (approximately the line from DC to Sacramento, CA).

The estimation shows increased urbanization (declining estimated IRR-values) throughout the entire continental US. On average, both for the northern and the southern counties, IRR drops at a decreasing rate as one moves from either coast towards the interior of the country. In the southern portion of the United States the changes are estimated to be negative (i.e., becoming more urban) throughout the entire range of longitudes. In the northern section, the estimated IRR changes approach zero in the midsection of the country, around 95 degrees western longitude, or around a the great circle just to the East of Fargo, ND. Note also that in the Eastern portion of the United States, urbanization is stronger in the southern US than in the northern US.

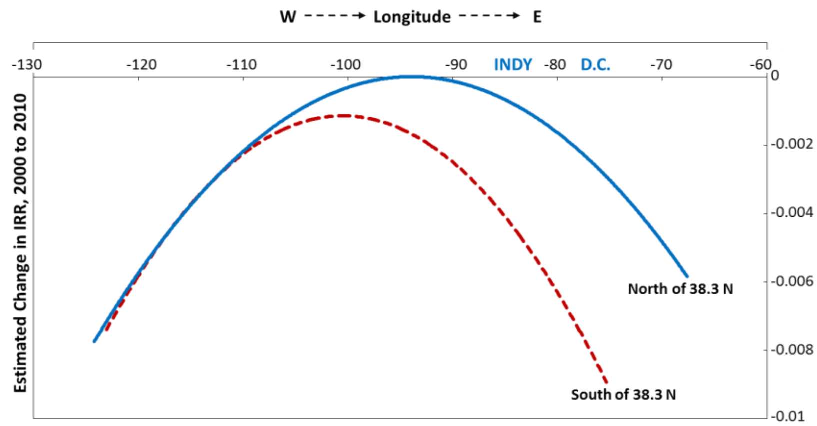


Figure 6. Estimated Rurality Change, $\Delta IRR\hat{}_{2000\text{-}10}$, as a Function of Longitude (AL and HI excluded)

5.4 The Use of IRR in a Regression Context

In applied and policy-oriented research, rurality is often considered an important variable contributing to variations in the outcome variable of interest. In a modelling context, rurality thus becomes a right-hand side variable. If rurality is measured discretely, it thus needs to be translated into dummy variables. Dummy variables are particularly valuable in the exploratory analysis, but their use is typically constrained to comparing means across categories. Moreover, very detailed discrete measures—such as RUCC and UIC—are avoided.[12] Instead, the most frequently used discrete rurality measure in regression models is the dummy variable metro.

If rurality is a continuous measure, such as the IRR, additional modelling options become available. Table 5 exemplifies how the index of relative rurality can be advantageously utilized when assessing rural-urban differences. As an outcome variable, the illustration uses counties' ethnic population compositions (% Hispanics).

On average, the percentage of Hispanics is significantly higher in metro counties than in nonmetro counties, with a metro/nonmetro gap of about one percentage point in 2010. Moreover, the variation in ethnic composition is quite substantial, both among metro and among nonmetro counties. Given that Hispanic settlement historically favored southern locations, a possible source of that variation is the location within the US. In a regression framework, such additional sources of variation can be taken into account.

In Table 5, we show the results of several regression models in which the share of Hispanics is regressed on rurality and location (variables *longitude*, *latitude*, and the dummy *Mexborder*=1 if the county is located in a state along the Mexican border (CA, AZ, NM or TX)). Models A to C measure rurality via the dummy variable *metro*. Model A suggests that, on average, Hispanic concentrations in metropolitan counties are 1.4 percentage points higher than in non-metropolitan counties. Models B and C suggest that the metro/nonmetro gap narrows substantially once the locational differentiation extends

[12] RUCC codes must be translated into 8, UIC codes into 11 dummy variables. In such situations, the dummies serve as fixed effects, and researchers rarely devote their attention to substantive interpretations of fixed effects.

beyond the southern border dummy. Nevertheless, they continue to predict that the populations of nonmetro counties have lower percentages of Hispanics than metropolitan county populations.

Table 5. Rurality Measures in Regression Models – An Example[a]

Dep. Var.: %Hispanics	**Discrete** Rurality Measure Metro Dummy			**Continuous** Rurality Measure Index of Relative Rurality		
n = 3,108	Model A	Model B	Model C	Model D	Model E	Model F
intercept	1.700	**-65.910**	2.561	**6.883**	**-65.763**	0.951
metro	**1.440**	**1.179**	**1.034**			
IRR				**-19.678**	**-71.046**	**-68.150**
IRR^2					**60.824**	**58.973**
latitude	**-0.392**	**-0.884**	**-1.645**	**-0.399**	*-1.068*	**-1.794**
latitude^2		**0.135**	**0.094**		**0.135**	**0.095**
longitude	**-0.202**	**-1.782**	**-0.657**	**-0.262**	**-2.251**	**-1.129**
longitude^2		**0.017**	**0.012**		**0.014**	**0.010**
*lat*long*		**0.114**	**0.067**		**0.112**	**0.067**
Mexborder	**22.641**		**13.277**	**22.178**		**12.888**
adj R^2	0.47	0.49	0.52	0.49	0.51	0.54

[a] Estimates in **bold** (*italic*) are significantly different from zero at the $\alpha = 0.01$ ($\alpha = 0.05$) level.

In Models D to F, rurality is measured using the IRR. The result of Model D is consistent with that of the Models A to C, suggesting that increasing rurality is associated with lower shares of Hispanics in the counties' populations. However, based on the expanded models E and F, the simple negative relationship between rurality and ethnic composition cannot be supported. Instead, we state that the relationship between ethnic composition and rurality is nonlinear. With increasing rurality, the share of the Hispanic population first declines but then increases again. Hispanic concentrations are high in urban counties and, to a lesser extent, in very rural counties. The minimum is estimated to be reached at a rurality level of approximately IRR = 0.58.

6. Conclusions

Rurality is a concept that is difficult to define and to measure. However, researchers and policy makers are in need of definitions and delineations of rural areas. Most approaches to defining rurality are thus pragmatic and, given our inability to define rurality, they cannot be properly validated.

In the United States, delineations / classifications / measurement of rurality and rural areas often utilize discrete measures based on thresholds. Most prominent is the rural-urban delineation of the U.S. Census Bureau. And, although not designed as a rural-urban classification, the metro/nonmetro differentiation of OMB has served as a proxy for a rural-urban differentiation in a good deal of social science research. Many other typologies have been developed, by researchers and government organizations alike. With a few exceptions, they are derivatives or extensions of the U.S. Census and OMB classifications.

The rural-urban classifications used in the United States share a reliance on thresholds and are thus subject to similar criticism: (1) thresholds are dependent on the set of spatial objects to be classified; (2) thresholds by design create artificial similarities and dissimilarities; (3) thresholds create discrete measures that are cumbersome to deal with in modelling frameworks. Moreover, in many cases the justifications for threshold choices are not clearly spelled out. However, one of the great advantages of the threshold-based rural-urban classifications is their simplicity. In particular, dichotomous classifications—such as the rural-urban classification of the U.S. Census Bureau—provide a clear answer to the "*What is rural*?" question that so often is of prime relevance for policy makers.

Continuous measures of rurality shifts the focus from the question of "*What is rural?*" to the more nuanced question of "*What is the degree of rurality*?" One example of a continuous measure is the Index of Relative Rurality. When designing a continuous measure like the IRR, choices need to be made about which dimensions to include, how to measure them, and how they should be linked. The IRR discussed in this paper includes dimensions frequently found in threshold based discrete measures: size, density, built-up area and proximity. They are transformed onto compatible scales and subsequently linked so that a score of *0* signals *extreme urban* and a score of *1* signals *extreme rural*. In the absence of theoretical guidance, the unweighted average was chosen as the most simple link function.

No matter which dimensions are included and how they are linked, continuous indices offer a number of advantages over discrete measures derived from a rural/urban classifications. First, rurality becomes a relative concept that allows us to rank areas by their degree of rurality and investigate their trajectories of rurality over time. This opens new avenues for understanding relationships between rurality and social issues. Second as continuous multi-dimensional measures, they are responsive to the multi-faceted nature of rurality, and sensitive to even small changes in one or several of the defining variables. Third, their design is scale independent, and they can be constructed for counties, groups of counties, and even very small scales such as townships. This is an important advantage over traditional classifications, particularly beneficial for designing and evaluating regional development strategies. Finally, continuous measures are easy to calculate, and analytically more easily handled than categories of a typology.

Despite these advantages of continuous measures of rurality, continuous measures of rurality are only a valuable addition to, but not a substitute for rural-urban classifications. In the policy realm, in particular, very nuanced measurement of rurality is inadequate in situations where decision making is dichotomous and requires a simple distinction between rural and urban.

References

Atav, A.S. and R. Darling 2012. Comparison of Coding Schemas and Birth Outcomes as Exemplars. *Online Journal of Rural Nursing and Health Care* 12(1), Spring 2012.

Ayres, J., B. Waldorf, M. McKendree and L. Hoelscher 2012. Defining Rural Indiana—The First Step. Purdue Center for Rural Development. *EC-766-W*. http://www. extension. purdue. edu/extmedia/EC/EC-766-W. pdf

Barber, C.E. 2013. Perceptions of Aging-Friendly Community Characteristics: Does County Rurality Make a Difference? *Online Journal of Rural Research & Policy* 8(2).

Berke, E.M., A.N. West, A.E. Wallace and W.B. Weeks 2009. Practical and Policy Implications of Using Different Rural-Urban Classification Systems: A Case Study of Inpatient Service Utilizarion Among Veterans Administration Users. *Journal of Rural Health* 25(3): 259-266.

Brezzi, M., L. Dijkstra and V. Ruiz 2011. OECD Extended Typology: The Economic Performance of Remote Rural Regions. *OECD Regional Development Working Papers* 2011/06, OECD Publishing. http://www.oecd.org/gov/regional-policy/48670214.pdf

Cromartie J. and S. Bucholtz 2008. Defining the "Rural" in Rural America. USDA ERS *Amber Waves* June 1, 2008.

De Montis, A., S. Caschili, and D. Trogu 2012. Accessibility, Rurality and Remoteness: An Investigation on the Island of Sardinia, Italy. In Planning Support Tools: Policy Analysis, Implementation and Evaluation. *Proceedings of the 7th International Conference on Informatics and Urban and Regional Planning INPUT 2012*: 155.

Domina, T. 2006. What Clean Break?: Education and Nonmetropolitan Migration Patterns, 1989–2004. *Rural Sociology* 71: 373–398.

Economic Research Service. Measuring Rurality. ERS Briefing http://www.ers.usda.gov/topics/rural-economy-population/rural-classifications.aspx

Fuguitt, G.V. 2005. Some Demographic Aspects of Rurality. *Research in Social Stratification and Mobility* 22: 73-90.

Gallardo, R. and R. Scammahorn 2012. Determinants of Innovative Versus Non-Innovative Entrepreneurs in Three Southern States. *The Review of Regional Studies* 41(2, 3): 103-117.

Ghelfi, L.M. and T. S. Parker 1997. A County-level Measure of Urban Influence. *Rural Development Perspectives* 12(2): 32-41.

Glasgow, N. and D.L. Brown 2012. Rural Ageing in the United States: Trends and Contexts. *Journal of Rural Studies* 28(4): 422-431.

Hubach, R.D., B. Dodge, T. Cola, P.R. Battani and M. Reece 2014. Assessing the Sexual Health Needs of Men Who Have Sex with Men (MSM) Residing in Rural and Mixed Rural Areas. *The Health Education Monograph Series* 31(2): 33-39.

Heflin, C. and M. Kathleen 2012. The Geography of Need: Identifying Human Service Needs in Rural America. *Journal of Family Social Work* 15(5): 359-374.

Ingram, D.D. and S.J. Franco 2014. 2013 NCHS Urban–Rural Classification Scheme for Counties. National Center for Health Statistics. Vital Health Stat 2(166). http://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf

Isserman, A.M. 2005. In the National Interest: Defining Rural and Urban Correctly in Research and Public Policy. *International Regional Science Review* 28(4): 465-499.

Johnson, K.M. and G.V. Fuguitt 2000. Continuity and Change in Rural Migration Patterns, 1950–1995. *Rural Sociology* 65: 27–49.

Kandel, W., J. Henderson, H. Koball and R. Capps 2011. Moving Up in Rural America: Economic Attainment of Nonmetro Latino Immigrants. *Rural Sociology* 76: 101–128.

Kaza, N. 2013. The Changing Urban Landscape of the Continental United States. *Landscape and Urban Planning* 110: 74-86.

Lambert, D.M., M.D. Wilcox, C.D. Clark, B. Murphy, and W.M. Park 2010. Is Growth in the Health Sector Correlated with Later-Life Migration? P*rogress in Spatial Analysis*. Springer Berlin Heidelberg: 381-403.

Levernier,W., M. D. Partridge, and D. S. Rickman 2000. The Causes of Regional Variations in U.S. Poverty: A Cross-county Analysis. *Journal of Regional Science* 40: 473-49.

Mammen, S., F.C. Lawrence, P.S. Marie, A.A. Berry, and S.E. Knight 2011. The Earned Income Tax Credit and Rural Families: Differences between Non-participants and Participants. *Journal of Family and Economic Issues* 32(3): 461-472.

Muhlenkamp, A, and B. Waldorf 2008. "Rural-urban Income Disparities among the Highly-educated." Chapter 8 in J. Poot, B. Waldorf and L. van Wissen (eds). *Migration and Human Capital*. Cheltenham, UK, Northampton, MA, USA, Edward Elgar.

OECD (2011). OECD Regional Typology. Directorate for Public Governance and Territorial Development. http://www.oecd.org/gov/regional-olicy/OECD_regional_typology_Nov2012.pdf

Office of Management and Budget 2000. Standards for Defining Metropolitan and Micropolitan Statistical Areas; Notice. Federal Register 65, No. 249, pp. 82228-82238.

Office of Management and Budget 2003. Revised Definitions of Metropolitan Statistical Areas, New Definitions of Micropolitan Statistical Areas and Combined Statistical Areas, and Guidance on Uses of the Statistical Definitions of These Areas. OMB Bulletin 03-04.

Pillarisetti, J.R. and J.C.J.M. van den Bergh 2010. Sustainable Nations: What Do Aggregate Indexes Tell Us? *Environment, Development and Sustainability* 12: 49-62.

Renkow, M. and D. Hoover 2000. Commuting, Migration, and Rural-Urban Population Dynamics. *Journal of Regional Science* 40(2): 261-287.

Stewart, L.A., and D.M. Lambert 2008. Factors Determining Corn-based Ethanol Plant Site Selection, 2000-2007. *Risk, Infrastructure and Industry Evolution.* Proceedings of a conference June 24-25, 2008, in Berkeley, California, pp. *69-78.*

United Nations Development Programme 2014. Human Development Report 2014 – Technical Notes. http://hdr.undp.org/sites/default/files/hdr14_technical_notes.pdf

Waldorf, B. 2006. A Continuous Multi-dimensional Measure of Rurality: Moving Beyond Threshold Measures. Selected Paper, Annual Meetings of the Association of

Agricultural Economics, Long Beach, CA, July 2006. http://ageconsearch.umn.edu/handle/21383

Waldorf, B. 2007a. Measuring Rurality. *INContext*. January 2007: pp. 5-8. http://www.incontext.indiana.edu/

Waldorf, B. 2007b. What is Rural and What is Urban in Indiana? Research Report, Purdue Center for Regional Development. PCRD-R-4. http://www.purdue.edu/dp/pcrd/pdf/PCRD-R-4.pdf

Waldorf, B., R.J.G.M. Florax and J. Beckhusen 2008. Spatial Sorting of Immigrants across Urban and Rural Areas in the United States: Changing Patterns of Human Capital Accumulation since the 1990s. *American Journal of Agricultural Economics* 90(5): 1312–18.

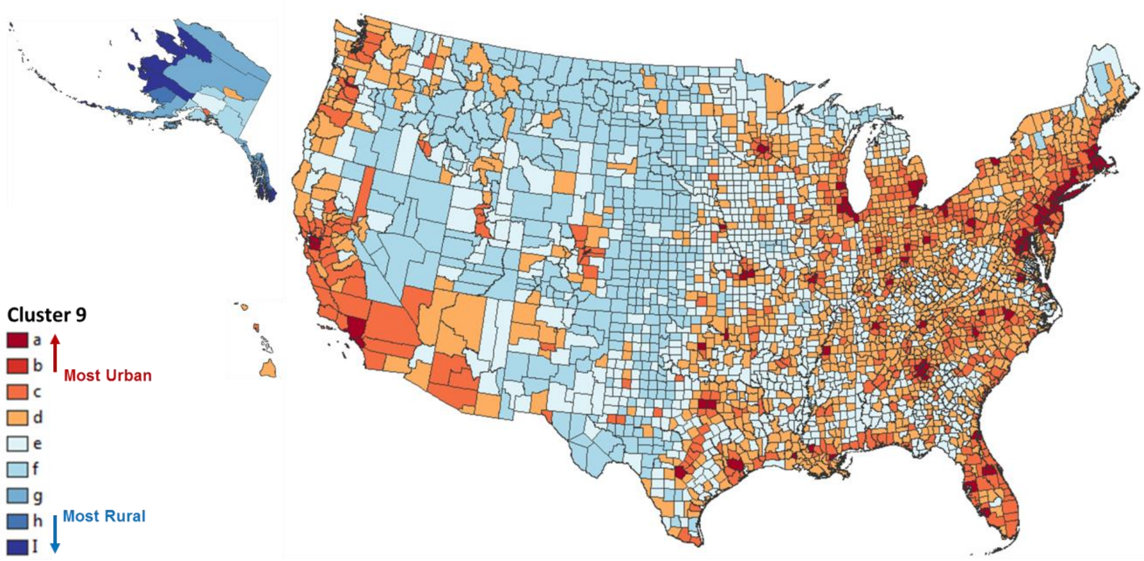
Weber, B.A. 2007: Rural Poverty: Why Should States Care and What Can State Policy Do? *Journal of Regional Analysis and Policy 37(1):48-52*.

Zhou, D., J. Xu. J. Radke, and L. Mu 2004. A Spatial Cluster Method Supported by GIS for Urban-Suburban-Rural Classification. *Chinese Geographical Science* 14(4): 337-342.

Appendix A.

Similarity-based typology of U.S. Counties 2010

9-cluster solution of a non-hierarchical cluster analysis with m = 4 variables:
– – population size, population density, % urban, and network distance – –



Note: Similarity-based typologies do not provide an *a priori* ranking by degree of rurality; here: ordering is based on IRR_{2010}.

Table A. Comparison of Classification by Cluster Analysis and IRR

Cluster (9)	#	Index of Relative Rurality 2010				
		Average	Stdev	Min	Max	Range
A	118	0.21	0.07	0.04	0.29	0.25
B	39	0.23	0.07	0.12	0.38	0.25
C	433	0.39	0.04	0.28	0.44	0.17
D	1,022	0.49	0.02	0.43	0.53	0.09
E	1,075	0.55	0.02	0.52	0.60	0.08
F	432	0.63	0.03	0.58	0.76	0.18
G	8	0.76	0.04	0.67	0.81	0.14
H	7	0.78	0.04	0.69	0.83	0.14
I	7	0.83	0.04	0.78	0.89	0.11
Grand Total	3,141	0.50	0.10	0.04	0.89	0.85

Appendix B:

Selection of Threshold-based Typologies in the U.S.

Classification System	Objects	Number of Types	Criteria									
			urban population [a]	OMB's CBSA	Contiguity	Size	Density	C-flows	Intrametro core-fringe distinction	Dist / Prox / Adj	Identity	IRR
Foundations												
US Census Bureau	blocks/tracts	2 (+1)			x	x	x					
OMB	counties	3	x		x	x		x				
Type-1												
Isserman 2005	counties	4	x			x	x					
VHA	points	3	x				x					
Type-2												
RUCC	counties	9	x	x		x				x		
NCES 2006 (Educ)	points	12	x	x		x				x		
HUD	places/counties	2	x	x		x						
Type-3												
Rural-Metropolitan Interface	counties	7		x		x			x	x		x
UIC	counties	12		x		x				x		
2013 NCHS	counties	6		x		x			x			
Type-4												
OECD (small units)	counties	2					x					
OECD (regions)	regions	3				x	x					
OECD (extended regions)	regions	5				x	x			x		
Ayres et al. 2012	counties	3				x	x				x	

[a] As defined by the U.S. Census Bureau

Appendix C:

IRR 2010 (top) and its Approximation via a 3rd - Order Polynomial Function of Latitude and Longitude